

STIC-Biotech/ChemLib

156005

NRFO

From: Swope, Sheridan
Sent: Thursday, June 09, 2005 4:16 PM
To: STIC-Biotech/ChemLib
Subject: 10/735,419

For 10/735,419, pls search and interference search:

SID 10: against the NT and AA data bases.

THANKS!!

Sheridan Swope, Ph.D.
Patent Examiner, AU 1652
Recombinant Enzymes
571-272-0943 (voice)
E02B71 Remsen Bld (Office)
E02C70 Remsen Bld (Mailbox)

No art

(724 ppg ID #10)

Rev. Swope

10-294aa

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2-_____
Date Searcher Picked up: _____
Date Completed: _____
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: _____ AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable
STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: _____
WWW/Internet: _____
Other(Specify): _____

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:57:41 ; Search time 174 Seconds

(without alignments)
865.237 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNINALVCGNGPSLKNID.....IRLPSDIKHVLYKEKANKR 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1571	100.0	294	2	Q7BP25	O7bp25 campylobact
2	823	52.4	291	2	Q9LAK3	Q9lak3 campylobact
3	816	51.9	291	2	Q938X6	Q938x6 campylobact
4	813	51.8	291	2	Q93CZ5	Q93cz5 campylobact
5	812	51.7	291	2	Q93MQ0	Q93mq0 campylobact
6	804	51.2	291	2	Q93D05	Q93d05 campylobact
7	804	51.2	291	2	Q9F0M9	Q9f0m9 campylobact
8	804	51.2	291	2	Q9L905	Q9l9q5 campylobact
9	690.5	43.3	430	2	Q9RGF1	Q9rgf1 campylobact
10	673.5	42.9	303	2	Q9CLP3	Q9clp3 pasteurella
11	441.5	28.1	231	1	Y352 HABIN	P24324 haemophilus
12	143	9.1	753	2	Q98S90	Q98s90 guillardia
13	137.5	8.8	232	2	Q7R8E8	O7rk88 plasmodium
14	137.5	8.8	3869	2	Q7R8H4	Q7rn84 plasmodium
15	136	8.7	754	2	Q7RGH4	O7rgi4 plasmodium
16	136	8.7	4095	2	Q6LFD1	Q6lfd1 plasmodium
17	135.5	8.6	1676	2	Q7RI39	O7ri39 plasmodium
18	133.5	8.5	974	2	Q8I304	Q8i304 plasmodium
19	133.5	8.5	1725	2	Q8IKU2	Q8iku2 plasmodium
20	132.5	8.4	1993	2	Q8ILC6	Q8ilc6 plasmodium
21	132	8.4	1847	2	Q8IDM0	Q8idm0 plasmodium
22	131	8.3	3787	2	Q8I1Z6	Q8i1z6 plasmodium
23	130.5	8.3	2120	2	Q7RQ47	O7rq47 plasmodium
24	130.5	8.3	2705	2	Q8I5L0	Q8i5l0 plasmodium
25	130	8.3	683	2	Q7RS77	Q7rs77 plasmodium
26	130	8.3	1679	2	Q8ILA0	Q8ila0 plasmodium
27	129.5	8.2	471	2	Q98RR3	Q98rr3 guillardia
28	129.5	8.2	850	2	Q6KID3	Q6kid3 mycoplasma
29	128.5	8.2	961	2	Q7RM52	O7rms2 plasmodium
30	128.5	8.2	2385	2	Q962I6	Q962i6 plasmodium
31	128.5	8.2	2400	2	Q8I3U7	Q8i3u7 plasmodium

RESULT 1

Q7BP25	PRELIMINARY;	PRT;	294 AA.
ID	Q7BP25	Q9PNF4;	
AC	Q7BP25; Q9PNF4;		
DT	05-JUL-2004 (TrEMBLrel. 27, Created)		
DT	05-JUL-2004 (TrEMBLrel. 27, Last sequence update)		
DT	25-OCT-2004 (TrEMBLrel. 28, Last annotation update)		
DE	Hypothetical protein Cj1140 (2,3-sialyl transferase) (Alpha-2,3-		
DE	sialyltransferase).		
GN	Name=cstIII; Synonyms=cst-III; OrderedLocusNames=Cj1140;		
OS	Campylobacter jejuni.		
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;		
OC	Campylobacteraceae; Campylobacter.		
OX	NCBI_TaxID=197;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MSC57360;		
RA	Guerry P., Ewing C.P., Moran A.P., Trust T.J.;		
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MSC 57360;		
RX	MEDLINE=20536403; PubMed=11083778;		
RX	DOI=10.1128/IAI.68.12.6656-6662.2000;		
RA	Guerry P., Ewing C.P., Hickey T.E., Prendergast M.M., Moran A.P.;		
RT	"Sialylation of lipooligosaccharide cores affects immunogenicity and serum resistance of Campylobacter jejuni.";		
RL	Infect. Immun. 68:6656-6662(2000).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=ATCC 43429; and ATCC 43429;		
RX	MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;		
RA	Gilbert M., Karwasaki M.-F., Bernatchez S., Young N.M., Taboada E.,		
RA	Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;		
RT	"The genetic basis for the variation in the lipo-oligosaccharide of the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated ganglioside mimics in the core oligosaccharide.";		
RL	J. Biol. Chem. 277:327-337(2002).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCTC 11168;		
RX	MEDLINE=20150912; PubMed=10688204; DOI=10.1038/35001088;		
RA	Farkhill J., Wren B.W., Mungall K.L., Kettle J.M., Churcher C.M.,		
RA	Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,		
RA	Jagels K., Kariyeh A.V., Moule S., Pallen M.J., Penn C.W.,		
RA	Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,		
RA	Whitehead S., Barrall B.G.;		
RT	"The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypervariable sequences.";		
RL	Nature 403:665-668(2000).		
DR	EMBL; AF195055; AAG29922.1; -		
DR	EMBL; AF257460; AAG37020.1; -		
DR	EMBL; AF400047; AAK85419.1; -		
DR	EMBL; AF044156; AAK73183.1; -		

ALIGNMENTS

32	128.5	8.2	5922	2	Q8IIN2	Q8iin2 plasmodium
33	128	8.1	389	2	Q6JPC2	Q6jpc2 neodiprion
34	128	8.1	402	2	Q8IL43	Q8il43 plasmodium
35	128	8.1	606	2	Q96277	Q96277 plasmodium
36	128	8.1	742	2	Q7REP2	Q7rep2 plasmodium
37	128	8.1	1391	2	Q7RA50	Q7ra50 plasmodium
38	128	8.1	1431	2	Q7RGT1	Q7rgt1 plasmodium
39	127.5	8.1	476	2	Q9EMJ9	Q9emj9 amsacta moo
40	127.5	8.1	1130	2	Q35596	Q35596 physarum po
41	127.5	8.1	1739	2	Q8IIJ5	Q8ijj5 plasmodium
42	127.5	8.1	1784	2	Q7RG50	Q7rg50 plasmodium
43	127	8.1	1083	2	Q8D2T9	Q8d2t9 wiggleswort
44	127	8.1	3182	2	Q7RI77	Q7ri77 plasmodium
45	127	8.1	5415	2	Q8IE74	Q8ie74 plasmodium

DR	EMBL; ALI39077; CAB73395.1; -
DR	PIR; G81318; G81318.
DR	GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
DR	InterPro; IPR009251; CST-I.
DR	Pfam; PF06002; CST-I; 1.
KW	Complete proteome; Glycosyltransferase; Hypothetical protein;
SW	SEQUENCE 294 AA; 35127 MW; 1729C9C35797D78F CRC64;
Query Match	
Best Local Similarity 100.0%; Score 1571; DB 2; Length 294;	
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 MSMMINALVCGSPSLKNIDYKRLPKQDFVRCNQFYFEDRYFGVKDVKYVFNFVFFFE 60
Db	1 MSMMINALVCGSPSLKNIDYKRLPKQDFVRCNQFYFEDRYFGVKDVKYVFNFVFFFE 60
Qy	61 QYTSKKLIQNEEYNIENIVCSTINLSEYIDGQFVDNPFELYSDFAFI GHEIIKKLKDFFA 120
Db	61 QYTSKKLIQNEEYNIENIVCSTINLSEYIDGQFVDNPFELYSDFAFI GHEIIKKLKDFFA 120
Qy	121 YIKNEYNRQRIITSGVYMCATAVALGVKSIYISGIDFYQDTNNLYAFDNNKKOLLNKCT 180
Db	121 YIKNEYNRQRIITSGVYMCATAVALGVKSIYISGIDFYQDTNNLYAFDNNKKOLLNKCT 180
Qy	181 GFKNQKPKFINHSMACDLQALDYLMKRYDNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db	181 GFKNQKPKFINHSMACDLQALDYLMKRYDNIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy	241 INDILIPDKYAERYGKSKRLKNLHYKLIKDLIRLPSDIKHVLEKRYANKNR 294
Db	241 INDILIPDKYAERYGKSKRLKNLHYKLIKDLIRLPSDIKHVLEKRYANKNR 294
RESULT 2	
Q9LAK3	PRELIMINARY; PRT; 291 AA.
ID	Q9LAK3
AC	Q9LAK3
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE	Alpha-2,3/8-sialyltransferase.
GN	Namescst-II;
OS	Campylobacter jejuni.
OC	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC	Campylobacteraceae; Campylobacter.
OX	NCBI_TaxID=197;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=OH4384;
RP	MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RX	Albert M., Brisson J.R., Karwaeki M.F., Michniewicz J.,
RA	Cunningham A.M., Wu Y., Young N.M., Wakarchuk W.W.;
RA	"Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT	Identification of the glycosyltransferase genes, enzymatic synthesis
RT	of model compounds, and characterization of nanomole amounts by 600-
RT	nmz (1)h and (13)c NMR analysis.";
RT	J. Biol. Chem. 275:3896-3906(2000).
RN	[2]
RN	SEQUENCE FROM N.A.
RP	STRAIN=ATCC 700297;
RC	Gilbert M., Michniewicz J., Wakarchuk W.W.;
RA	Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL	EMBL; AF130984; AAF31771.1; -
DR	ENBL; AF216647; AAF36462.1; -
DR	PDB; 1R07; X-ray; A/B/C/D=1-259.
DR	PDB; 1R08; X-ray; A/B=1-259.
DR	GO; GO:0016757; F:transferase activity, transferring glycosyl. .; IEA.
DR	InterPro; IPR009251; CST-I.
DR	Pfam; PF06002; CST-I; 1.
SW	Glycosyltransferase; Transferase.
SW	SEQUENCE 291 AA; 34544 MW; AB0741D5FCEDAB6A CRC64;


```

Db 184 HYHGSKNTDIIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy 244 ILIPDKYAQERYG-----KKSRLKENLHYKLIKIDLRPSDIKHLYKEK 288
Db 243 ILIPSSA-----YGFKFTKNINFKKIKIKENIYKLIKIDLLRPSDIKHLYFKGK 291

RESULT 4
Q93CZ5 PRELIMINARY; PRT; 291 AA.
AC Q93CZ5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Bifunctional alpha-2,3-/2,8-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43449;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AF400048; AAK91725.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34504 MW; 50162DB8D81559F CRC64;

Query Match 51.8%; Score 813; DB 2; Length 291;
Best Local Similarity 54.3%; Pred. No. 9.9e-49;
Matches 159; Conservative 43; Mismatches 73; Indels 18; Gaps 6;

Qy 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFVEDRYFVGDKVYFNFVFFPEQYTSKK 67
Db 5 IISGNPSLKEIDYSLRPNDFVRCNQFVEDRYFVGDKVYFNFVFFPEQYTSKK 64
Qy 68 LIQNEEYNIENIVCGTINLEYIDGQFVDNFELYSDFDAFLGHEIHKLDKDFAYIKYNEI 127
Db 65 LIQNEEYIEILMCSNYNQAHLENENFVKTFYDFPDHGLGYDFKQKFEFNAYFKFHEI 124
Qy 128 YNQRTISGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLKCTGFKNQKF 187
Db 125 YFNQRTISGVYMCATVALGYKSIYISGIDFY-DNGGGYAFDTKQKLLKLPNFKNDNS 183
Qy 188 KFTNHSMACDLQALDYLKMYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
Db 184 HYHGSKNTDIIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy 244 ILIPDKYAQERYG-----KKSRLKENLHYKLIKIDLRPSDIKHLYKEK 288
Db 243 ILIPSSA-----YGFKFTKNINFKKIKIKENIYKLIKIDLLRPSDIKHLYFKGK 291

RESULT 5
Q93MQ0 PRELIMINARY; PRT; 291 AA.
AC Q93MQ0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3-/alpha-2,8-sialyltransferase.
GN Name=cstII;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;

```

```

OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43438;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AF400048; AAK91725.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34572 MW; BD1BEDEF6F521E6 CRC64;

Query Match 51.7%; Score 812; DB 2; Length 291;
Best Local Similarity 54.3%; Pred. No. 1.2e-48;
Matches 159; Conservative 46; Mismatches 70; Indels 18; Gaps 6;

Qy 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFVEDRYFVGDKVYFNFVFFPEQYTSKK 67
Db 5 IISGNPSLKEIDYSLRPNDFVRCNQFVEDRYFVGDKVYFNFVFFPEQYTSKK 64
Qy 68 LIQNEEYNIENIVCGTINLEYIDGQFVDNFELYSDFDAFLGHEIHKLDKDFAYIKYNEI 127
Db 65 LIQNEEYETELIMCSNYNQAHLENENFVKTFYDFPDHGLGYDFKQKFEFNAYFKFHEI 124
Qy 128 YNQRTISGVYMCATVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLKCTGFKNQKF 187
Db 125 YLNQRTISGVYMCATVALGYKSIYISGIDFYQNGSS--YAFDTKQENLLKLPDFFKNDRS 183
Qy 188 KFTNHSMACDLQALDYLKMYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKPKKYIND 243
Db 184 HYHGSKNTDIIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy 244 ILIPDKYAQERYG-----KKSRLKENLHYKLIKIDLRPSDIKHLYKEK 288
Db 243 ILIPSSA-----YGFKFTKNINFKKIKIKENIYKLIKIDLLRPSDIKHLYFKGK 291

RESULT 6
Q93D05 PRELIMINARY; PRT; 291 AA.
AC Q93D05;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3-sialyltransferase.
GN Name=cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43456;
RX MEDLINE=21634908; PubMed=11689567; DOI=10.1074/jbc.M108452200;
RA Gilbert M., Karwaski M.-F., Bernatchez S., Young N.M., Taboada E.,
RA Michniewicz J., Cunningham A.-M., Wakarchuk W.W.;
RT "The genetic basis for the variation in the lipo-oligosaccharide of
RT the mucosal pathogen, Campylobacter jejuni: Biosynthesis of sialylated
RL J. Biol. Chem. 277:327-337(2002).
DR EMBL; AF401528; AAL05990.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34640 MW; F294A04ACBA82882 CRC64;

```

```
Query Match      51.2%; Score 804; DB 2; Length 291;
Best Local Similarity 53.9%; Pred. No. 4.2e-48;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGDKVYVFNPFVFFQYVTSKK 67
DB 5 IIAAGNPSLKEIDYSRLPNDPFRVRCNQFYFEDRYFVGDKVYVFNPFVFFQYVTSKK 64
QY 68 LIQNEEYNIENIVCSTINLEIDYDGFQVDFNPFYFSDAFGLGHEIIKKLDPFAYIKYNEI 127
DB 65 LIQNEEYTELIMCSNNOAHLENENFVKTFYDYPDAHLGDFYFQKLEFNAYFKFHEI 124
QY 128 YNRQRITSGVYMCATVALGYKSIYISGIDFYDTNNLYAFDNNKNLLNCKTGCFKNQKF 187
DB 125 YFNQRITSGVYMCATVALGYKSIYISGIDFYDTNNLYAFDNNKNLLNCKTGCFKNQKF 183
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKKPKKYIND 243
DB 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLPCNSLLANFIELAPLNSNFIQEK--NNYTKD 242
QY 244 ILIPDKVAQERYG-----KKSRLENLHYKLIKDLRLPSDIKHLYKEK 288
DB 243 ILIPSSSEA-----YGFKSKNINFKKIKENYVYKLIKDLRLPSDIKHLYKEK 291

RESULT 7
ID Q9F0M9 PRELIMINARY; PRT; 291 AA.
AC Q9F0M9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2.3-sialyltransferase.
GN Name-cst-II;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43432;
RA Gilbert M., Michniewicz J., Wakarchuk W.W.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF215659; AAG43979.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34553 MW; E41B594ACD7280F8 CRC64;

Query Match      51.2%; Score 804; DB 2; Length 291;
Best Local Similarity 53.9%; Pred. No. 4.2e-48;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGDKVYVFNPFVFFQYVTSKK 67
DB 5 IIAAGNPSLKEIDYSRLPNDPFRVRCNQFYFEDRYFVGDKVYVFNPFVFFQYVTSKK 64
QY 68 LIQNEEYNIENIVCSTINLEIDYDGFQVDFNPFYFSDAFGLGHEIIKKLDPFAYIKYNEI 127
DB 65 LIQNEEYTELIMCSNNOAHLENENFVKTFYDYPDAHLGDFYFQKLEFNAYFKFHEI 124
QY 128 YNRQRITSGVYMCATVALGYKSIYISGIDFYDTNNLYAFDNNKNLLNCKTGCFKNQKF 187
DB 125 YFNQRITSGVYMCATVALGYKSIYISGIDFYDTNNLYAFDNNKNLLNCKTGCFKNQKF 183
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDFVLSKKPKKYIND 243
DB 184 HYIGHSKNTDIKALEFLEKTYKIKLYCLPCNSLLANFIELAPLNSNFIQEK--NNYTKD 242
QY 244 ILIPDKVAQERYG-----KKSRLENLHYKLIKDLRLPSDIKHLYKEK 288
DB 243 ILIPSSSEA-----YGFKSKNINFKKIKENYVYKLIKDLRLPSDIKHLYKEK 291
```

```
Db 243 ILIPSSSEA-----YGFKSKNINFKKIKENYVYKLIKDLRLPSDIKHLYKEK 291

RESULT 8
Q9L9Q5 PRELIMINARY; PRT; 291 AA.
AC Q9L9Q5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Alpha-2.3-sialyltransferase.
GN Name-cst-II; Synonyms=cst;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43446;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brisson J.R., Karwaski M.F., Michniewicz J.,
RA Cunningham A.M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT mhz (1)h and (13)c NMR analysis.";
RT J. Biol. Chem. 275:3896-3906(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43446;
RA Gilbert M.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=81-176;
RX MEDLINE=21655167; PubMed=11796612; DOI=10.1128/IAI.70.2.787-793.2002;
RA Guerry P., Szymanski C.M., Prendergast M.M., Hickey T.B., Ewing C.P.,
RA Pattarini D.L., Moran A.P.;
RT "Phase variation of Campylobacter jejuni 81-176 lipooligosaccharide
RT affects ganglioside mimicry and invasiveness in vitro.";
RL Infect. Immun. 70:787-793(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=GB11;
RX PubMed=14742567; DOI=10.1128/IAI.72.2.1162-1165.2004;
RA Gilbert M., Godschalk P.C., Karwaski M.F., Ang C.W., Van Belkum A.,
RA Li J., Wakarchuk W.W., Endtz H.P.;
RT "Evidence for Acquisition of the Lipooligosaccharide Biosynthesis
RT Locus in Campylobacter jejuni GB11, a Strain Isolated from a Patient
RT with Guillain-Barre Syndrome, by Horizontal Exchange.";
RL Infect. Immun. 72:1162-1165(2004).
DR EMBL; AF167344; AAF34137.1; -.
DR EMBL; AF305571; AAL09368.1; -.
DR EMBL; AY422197; AAR82875.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . .; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
DR Glycosyltransferase; Transferase.
SQ SEQUENCE 291 AA; 34610 MW; 1BE4A04ACD728228 CRC64;

Query Match      51.2%; Score 804; DB 2; Length 291;
Best Local Similarity 53.9%; Pred. No. 4.2e-48;
Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGDKVYVFNPFVFFQYVTSKK 67
DB 5 IIAAGNPSLKEIDYSRLPNDPFRVRCNQFYFEDRYFVGDKVYVFNPFVFFQYVTSKK 64
QY 68 LIQNEEYNIENIVCSTINLEIDYDGFQVDFNPFYFSDAFGLGHEIIKKLDPFAYIKYNEI 127
DB 65 LIQNEEYTELIMCSNNOAHLENENFVKTFYDYPDAHLGDFYFQKLEFNAYFKFHEI 124
QY 128 YNRQRITSGVYMCATVALGYKSIYISGIDFYDTNNLYAFDNNKNLLNCKTGCFKNQKF 187
```

```
Db 125 YFNQRTSGVMCAVAIAAGYKIYSGIDFYQNGSS-YAFDTKQENLLKLAPDFNDRS 183
QY 188 KFINHSMACDQALDYLKMDYDNIYSL--NS--DEYFKLAPDYGSDVLSKPKKYIND 243
Db 184 HYIGHSKNTDKALEFLEKTYKIKLYCLPNSLLANFELAPLNSGFIQEK-NNYTKD 242
QY 244 ILIPDKYAQRYYG-----KKSRLKENLHVKLIKOLIRLSPDIKHLYKEK 288
Db 243 ILIPSSSEA----YKFSKINFFKIKIKENVYKLIKOLLRLSPDINGYFKGK 291

RESULT 9
Q9RGF1 PRELIMINARY; PRT; 430 AA.
AC Q9RGF1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha-2,3-sialyltransferase.
GN Name-cst-I;
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Campylobacteraceae; Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OH4384;
RX MEDLINE=20127862; PubMed=10660542; DOI=10.1074/jbc.275.6.3896;
RA Gilbert M., Brieson J.R., Karwasaki M.F., Michniewicz J.,
RA Cunningham A.M., Wu Y., Young N.M., Wakarchuk W.W.;
RT "Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
RT Identification of the glycosyltransferase genes, enzymatic synthesis
RT of model compounds, and characterization of nanomole amounts by 600-
RT mhz (1)h and (13)c NMR analysis.";
RL J. Biol. Chem. 275:3896-3906(2000).
DR EMBL; AF130466; AAF13495.1; -.
DR GO; GO:0016757; F:transferase activity, transferring glycosyl. . . ; IEA.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Glycosyltransferase; Transferase.
SQ SEQUENCE 430 AA; 50396 MW; 1D03B6797169425C CRC64;

Query Match 43.3%; Score 680.5; DB 2; Length 430;
Best Local Similarity 49.7%; Pred. No. 2.4e-39;
Matches 145; Conservative 49; Mismatches 75; Indels 23; Gaps 8;

QY 1 MSNII-NALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVKYVFPNPFVF 59
Db 12 VSKNMQNIITAGNGPSLKNINYLKRLPREYDVRCNQYFEDRYFVGKDVKYVFPNPFVF 71
QY 60 EQYTSKKLQNEBYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHEIILKLDKDF 119
Db 72 QQYHTAKQLIKNEYIKNIFCSTENLPFIESNDFLHQFYFPDPAKLGVEVIENLKEFY 131
QY 120 AYIKNEYINRQITSGVVMCAVALGYKSIYISGIDFYQDTNNLYAFDNKNKLLNKC 179
Db 132 AYIKNEYINRQITSGVVMCAVALGYKSIYISGIDFYQDTNNLYAFDNKNKLLNKC 179
QY 180 TGFKNQKFEIN-HSMACDQALDYLKMDYDNIYSLNSD-----EYFKLAPDYGSDV 234
Db 191 PGIKD--FKPSNCHSKYDEALUKLSYIKVNIYALCDDSIANHPPLSINNNFTLE 248
QY 235 KPKPKYINDILIPDKYAQRYYGKKGRLKEN-----LHYK--LIKDL 274
Db 249 NKHNSINDILLTNTDPGVVSFY--KNQLKADKNIMLNFNLYLHSDKNLIKFL 298

RESULT 10
Q9CLP3 PRELIMINARY; PRT; 303 AA.
AC Q9CLP3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
```

```
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein PM1174.
GN OrderedLocusNames=PM1174;
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100; DOI=10.1073/pnas.051634598;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RT "Complete genomic sequence of Pasteurella multocida Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
DR EMBL; AE006157; AAK03258.1; -.
DR InterPro; IPR009251; CST-I.
DR Pfam; PF06002; CST-I; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 303 AA; 35676 MW; BA7B8C87563B921 CRC64;

Query Match 42.9%; Score 673.5; DB 2; Length 303;
Best Local Similarity 46.8%; Pred. No. 5.1e-39;
Matches 137; Conservative 47; Mismatches 94; Indels 15; Gaps 7;

QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVKYVFPNPFVFPEQYTSKK 67
Db 14 IVAGNGESLSQIDYRLPKNDYDRCNQYFERYFLGNKIKAVFPTGCVLEQYTYLYH 73
QY 68 LIQNEBYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHE-IILKLDKDFAYIKYNE 126
Db 74 LKRNNEYVDNVLSSFNHPTVD-LEKSQIKALFIDVINGEYKLSKLTAFDVIYLYKE 132
QY 127 IYNRQITSGVVMCAVALGYKSIYISGIDFYQDTNNLYAFDNKNKLLNKC TGPKNQK 186
Db 133 LYENRQITSGVVMCAVAIAAGYTDIVLTGIDFYQASEENYAFDNKKPIRLLPDPKKE 192
QY 187 FFIHNSMACDQALDYLKMDYDNIYSLNS-----DEYFKLAPDYGSD-----FVLSKKPK 238
Db 193 TLFYSYHSDKIDLEALSFLQHYVNFYSISPSPLSKFPPI-PTVEDDCETTFVAPLK-B 250
QY 239 KYINDILIPDKYAQR--YYGKSKRLKENLHVKLIKOLIRLSPDIKHLYKEK 288
Db 251 NYINDILLPFPVYEKLGITVSKSRFNSLIVRLIRDLLKLPALKHYLKEK 303

RESULT 11
Y352_HAEIN STANDARD; PRT; 231 AA.
ID Y352_HAEIN
AC F24324;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein HI0352 (ORF1).
GN OrderedLocusNames=HI0352;
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RM 7004 / Serotype B;
RX MEDLINE=92065797; PubMed=1956282;
RA Maskell D.J., Szabo M.J., Butler P.D., Williams A.E., Moxon E.R.;
RT "Molecular analysis of a complex locus from Haemophilus influenzae
RT involved in phase-variable lipopolysaccharide biosynthesis.";
RL Mol. Microbiol. 5:1013-1022(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=RD / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
```



```
Db 1652 LSNLDGRKSEONKV---NIFHKOGNLFYQRRKYQREGANNIVTFPAKEKKGKSDVSNIL 1708
Qy 64 TSKLQIONEEYNIENI-----VCSTINLEYIDGQFVDNPFELYSFSDAPLGH 109
Db 1709 STNKLNIHKNSNDSIDSYEKIEKVRSGIINTPSKRYI-----HREFLSDSPETN 1760
Qy 110 EILKLLK--DFFAVIKYNE---IYNQRITSGVVMCATAVAGYKSIYISGIDPYQDTNN 164
Db 1761 SFRKNKSSETIQIPKPRORTIRNN-----SDHMGISYDDKFI-----EKN 1805
Qy 165 LYAFDNNKQNLNCKTGFKN-----QKFKPIN-HSMACDLQALDYLKMKRYDVIYSLNSD 218
Db 1806 INRYSEKSK-----FKSNLKKKSNYVDIRSIKSKSTLEYFVEK-----D 1847
Qy 219 EYFKLAPDIGSD-FVLKSKPKKYINDILIPDKYAQERYGKSRLENLHYKL--INKLI 275
Db 1848 YLKQWTTSTISDSYEIKDKTKYKKNKTIPKKNKKEKIEKKKXKKNKKEKKKKDIL 1907
Qy 276 RLPSDIKHYLKEKYANKN 293
Db 1908 IYPOIKSYVKEK--INKN 1924

RESULT 14
Q7RN84 PRELIMINARY; PRT; 3869 AA.
AC Q7RN84 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE Hypothetical protein.
GN Name=PY01937;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73233;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteza M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL0100527; EAA21339.1; -.
DR InterPro; IPR002917; MMR_HSR1; 1.
DR Pfam; PF01926; MMR_HSR1; 1.
SQ SEQUENCE 3869 AA; 457791 MW; E0DF5A5603674201 CRC64;

Query Match 8.8%; Score 137.5; DB 2; Length 3869;
Best Local Similarity 22.5%; Pred. No. 1.4; Indels 127; Gaps 20;
Matches 73; Conservative 48; Mismatches 77;

Qy 27 QFDV-----FRCNQFYEDRYFGVKGVKVFNFVFPFQYTSKLIQNEEYN 75
Db 3436 QYDVLHGCLLYINPNNQL---DHEFFIAKEIPLVLF-----LIQNKQVS 3478
Qy 76 -----IENIVCSTINLEYIDGQFVDNPFELYSFSDAPLGH 108
Db 3479 RDIKFGENVNRTNISNRCVCDVNYTELKNVCNCINIDIKNEFINNIE----- 3529
Qy 109 HEIILKLDKDFAYIKYN-----EIVN--RQRTISGVVMCATAVAGYKSIYISGIDFY 159
```

```
Db 3530 -EKIPKYNVYIEYNNMWMKMYEIEAIEBKIS--YQYNNRYKYNKKVAYFKGFNIY 3586
Qy 160 QDT-----NNLYAFDNNKQNLNCKTGFKNQKPKFINHSMACDLQALDYLKMKRYDYN 211
Db 3587 EKENYEQEIKFNNYCIYTVYKKKNIKEKNIYNN-----YTV-YNNGN 3627
Qy 212 IYSLNSDEYFKLAPDIGSDPVLKSKPKKYINDILIPDKYAQERYGKSRLENLHYKLI 271
Db 3628 YLVNTE-----DV-LDF--EKNPETYVN--VEEPY---KKFEKISLNLENI--KIE 3670
Qy 272 KD---LIRLPSDIKHYLKEKYANKN 293
Db 3671 NEGFCVLVSLKNE--NNEVK---GNKN 3691

RESULT 15
Q7RG14 PRELIMINARY; PRT; 754 AA.
AC Q7RG14 (TREMBlrel. 26, Created)
DT 01-MAR-2004 (TREMBlrel. 26, Last sequence update)
DE Probable trRNA modification GTPase time.
GN Name=PY04363;
OS Plasmodium yoelii yoelii.
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=73239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=17XNL;
RX PubMed=12368865; DOI=10.1038/nature01099;
RA Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Perteza M.,
RA Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA Cho J.K., Quackenbush J., Sedegah M., Shoabi A., Cummings L.M.,
RA Florens L., Yates F.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA Carucci D.J.;
RT "Genome sequence and comparative analysis of the model rodent malaria
RT parasite Plasmodium yoelii yoelii."
RL Nature 419:512-519(2002).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABL01001318; EAA16213.1; -.
DR InterPro; IPR002917; MMR_HSR1; 1.
DR Pfam; PF01926; MMR_HSR1; 1.
SQ SEQUENCE 754 AA; 88639 MW; 2826844083775506 CRC64;

Query Match 8.7%; Score 136; DB 2; Length 754;
Best Local Similarity 26.0%; Pred. No. 0.3;
Matches 81; Conservative 35; Mismatches 85; Indels 110; Gaps 19;

Qy 17 KNIDYKRL--PKQFD-----VFRCNQFYED-RYFVGKDVK--YVFNFVFPFQYTSK 66
Db 211 KNIESRKLKYKIYDINSNDIIDNVYAFKSPNSYTGEDVVEIYCHGNPFI----- 262
Qy 67 KLIQNEEYNIENIVCSTINLEYIDGQFVDNPFELYSFSDAPLGHIEIK---KLKDFPAYIK 123
Db 263 KEIMNAIDHNNIMYIINDE-----YNNDDWYFNQIEQNIQEQIETIDLNNFVIRSKKG 301
Qy 124 YNEIYNQRITSGVVMCATAVAGYKS---IYISGIDFYQ-----DTNN----- 164
Db 302 HDENVN-----YNNDDWYFNQIEQNIQEQIETIDLNNFVIRSKKG 343
Qy 165 ---LYAFDNNKQNLNCKTGFKNQKPKFINHSMACDLQ-----ALDYLKMKRYDVIY--SL 215
Db 344 EFTIRAFENNNQWLL-QIEGLKELLF-----CKQIKQKIALNYL--NGYAKNIVLKL 393
Qy 216 NSD-----EYFKLAPDIGSDPVLKSKPKKYINDILIPDKYAQERYGKSRLENLHYKL 270
```

Db	394	RNDIKLLIYIQLKIDFEDDHIITKGGKYIN-----MYIKKKVNNNSIKKHKEILKRN	447
Qy	271	IKDLIRLPDI	281
Db	448	IESL-NTPSNV	457

Search completed: June 13, 2005, 21:09:23
Job time : 177 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 20:53:46 ; Search time 161 Seconds
(without alignments)
706.258 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNINALVCGNPSLKNID.....IRLPDIKHVLYKEVANKNR 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	294	3	AAY97215 Campyloba
2	1571	100.0	294	6	ABJ18483 Campyloba
3	823	52.4	291	3	AAY97204 Campyloba
4	823	52.4	291	6	ABJ18479 Campyloba
5	817	52.0	291	3	AAY97211 Campyloba
6	817	52.0	291	6	ABJ18481 Campyloba
7	813	51.8	291	3	AAY97210 Campyloba
8	813	51.8	291	6	ABJ18480 Campyloba
9	804	51.2	291	3	AAY97212 Campyloba
10	804	51.2	291	6	ABJ18485 Campyloba
11	804	51.2	291	6	ABJ18484 Campyloba
12	804	51.2	291	6	ABJ18482 Campyloba
13	680.5	43.3	430	2	AAY45221 Campyloba
14	673.5	42.9	303	6	ABR40195 Campyloba
15	128.5	8.2	2013	3	AAB18265 Plasmodiu
16	124.5	7.9	1398	3	AAB18292 Plasmodiu
17	123	7.8	761	3	AAB16529 Bacteriop
18	121.5	7.7	1182	3	AAB18288 Plasmodiu
19	119	7.6	225	6	ABU25016 Protein e
20	116.5	7.4	353	2	AAY19859 B. burgdo
21	116.5	7.4	373	3	AAY19858 Plasmodiu
22	116.5	7.4	1817	3	AAB18301 Plasmodiu
23	116	7.4	726	5	ABB77613 AmEPV NTP
24	115	7.3	295	5	ABB09489 AmEPV fir
25	115	7.3	1121	3	AAB18241 Plasmodiu

ALIGNMENTS

RESULT 1

AAY97215

ID AAY97215 standard; protein; 294 AA.

XX AAY97215;

XX 12-SEP-2003 (revised)

DT 22-DEC-2000 (first entry)

XX

DE Campylobacter jejuni NCTC11168 alpha-2,3-sialyltransferase.

XX

KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;

KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;

KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;

KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;

KW immunity; immunogen; ganglioside.

XX

OS Campylobacter jejuni; NCTC11168.

XX

PN WO200046379-A1.

XX

PD 10-AUG-2000.

XX

PF 01-FEB-2000; 2000WO-CA000086.

XX

PR 01-FEB-1999; 99US-0118213P.

XX

PR 31-JAN-2000; 2000US-00495406.

XX

PA (CANA) NAT RES COUNCIL CANADA.

XX

PI Gilbert M, Wakarchuk WW;

XX

XX WPI; 2000-524418/47.

XX

PT Novel glycosyltransferase polypeptides and polynucleotides useful for

PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic

PT reagents and as immunogen for producing antibodies.

XX

PS Disclosure; Page 97-98; 120pp; English.

XX

CC A reaction mixture for the synthesis of a sialylated oligosaccharide is

CC useful for synthesising sialylated oligosaccharides such as ganglioside,

CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for

CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and

CC other oligosaccharides that have biological activity. The enzymes and

CC nucleic acids that encode them are useful for studies of the pathogenesis

CC mechanisms of organisms that synthesize ganglioside mimics, such as C.

CC jejuni and the nucleic acids are used as probes to study expression of

CC

CC

CC

CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analyzing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the Campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing
 CC reagents or as therapeutics and as immunogens for producing antibodies.
 CC Bacterial glycosyltransferase can be used to catalyse the formation of
 CC oligosaccharides that are identical to the corresponding mammalian
 CC structures and are easier and less expensive to produce in large
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial
 CC origin of the enzymes facilitates expression of large quantities of the
 CC enzymes using relatively inexpensive prokaryotic expression systems.
 CC (Updated on 12-SEP-2003 to standardise OS field)
 XX
 SQ Sequence 294 AA;
 Query Match 100.0%; Score 1571; DB 3; Length 294;
 Best Local Similarity 100.0%; Pred. No. 1e-135;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSNMNINAVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
 DB 1 MSNMNINAVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
 QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIHKLDKFFA 120
 DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIHKLDKFFA 120
 QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 QY 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
 DB 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
 QY 241 INDILIPDKYAQERYYGKSKRLKENLHYKLIKOLIRLPSDIKHYLKEKYANKNR 294
 DB 241 INDILIPDKYAQERYYGKSKRLKENLHYKLIKOLIRLPSDIKHYLKEKYANKNR 294
 RESULT 2
 ABJ18483
 ID ABJ18483 standard; protein; 294 AA.
 XX
 AC ABJ18483;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 XX Campylobacter jejuni bifunctional sialtransferase catII #7.
 DE
 XX Enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.
 XX
 OS Campylobacter jejuni.
 XX
 XX WO200274942-A2.
 XX
 XX 26-SEP-2002.
 PD
 XX 22-FEB-2002; 2002WO-CR0000229.
 PF
 XX 21-MAR-2001; 2001US-00816028.
 XX
 XX (CANADA) NAT RES COUNCIL CANADA.
 PA
 XX Gilbert M, Wakarchuk WW;
 FI

XX WPI; 2003-040554/03.
 DR
 XX New glycosyltransferases from Campylobacter, useful for synthesizing
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 PT mechanisms of organisms that synthesize ganglioside mimetics.
 XX
 PS Claim 5; Page 98; 107pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
 CC Acetylgalactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of Campylobacter enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. jejuni oligosaccharides of the invention may be used as
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
 CC metastasis). The present amino acid sequence represents a Campylobacter
 CC jejuni protein of the invention
 XX
 SQ Sequence 294 AA;
 Query Match 100.0%; Score 1571; DB 6; Length 294;
 Best Local Similarity 100.0%; Pred. No. 1e-135;
 Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MSNMNINAVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
 DB 1 MSNMNINAVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
 QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIHKLDKFFA 120
 DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIHKLDKFFA 120
 QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 QY 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
 DB 181 GFKNQKFKFINHSMACDLQALDYLKRYDNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
 QY 241 INDILIPDKYAQERYYGKSKRLKENLHYKLIKOLIRLPSDIKHYLKEKYANKNR 294
 DB 241 INDILIPDKYAQERYYGKSKRLKENLHYKLIKOLIRLPSDIKHYLKEKYANKNR 294
 RESULT 3
 AAY97204
 ID AAY97204 standard; protein; 291 AA.
 XX
 AC AAY97204;
 XX
 DT 12-SEP-2003 (revised)
 DT 22-DEC-2000 (first entry)
 XX
 XX Campylobacter jejuni OH4384 CstII sialyltransferase.
 XX
 XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
 KW immunity; immunogen; ganglioside.
 XX
 OS Campylobacter jejuni; OH4384.
 XX
 XX WO2000046379-A1.
 XX
 XX 10-AUG-2000.
 PD


```

CC Acetyl(galactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present amino acid sequence represents a Campylobacter
CC jejuni protein of the invention
XX
SQ Sequence 291 AA;

Query Match 52.0%; Score 817; DB 6; Length 291;
Best Local Similarity 54.3%; Pred. No. 1.6e-66;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 67
DB 5 ILAGNGPSLKEIDYSLRPNDVDFRCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 64
QY 68 LIQNEEYENIENIVCSTINLEYIDGQFVDNFELFSDAFILGHEIIKKLDFPFIYKNEI 127
DB 65 LIQNEEYETELIMCSNFNQAHLENQNFVKTIFYDFPDALHGYDFPKQKEFNAYFKFHEI 124
QY 128 YNQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNKCTGFKNQKF 187
DB 125 YFNQRITSGVYMCATAVALGYSIYISGIDFYQNGSS-YAFDTKQKLLKLPNFKNDNS 183
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDPVLSPKPKYIND 243
DB 184 HYIGHSKNTDIALKLEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK-NNYTKD 242
QY 244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLRLPSDIKHLYLKEK 288
DB 243 ILIPSSSEA----YGFKSNINFKIKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 7
AA97210
ID AA97210 standard; protein; 291 AA.
XX
AC AA97210;
XX
XX 12-SEP-2003 (revised)
XX 22-DEC-2000 (first entry)
XX
DE Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.
XX
KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside.
XX
OS Campylobacter jejuni; O:10 serotype.
XX
XX WO200046379-A1.
XX
XX 10-AUG-2000.
XX
XX 01-FEB-2000; 2000WO-CA000086.
XX
XX 01-FEB-1999; 99US-0118213P.
XX 31-JAN-2000; 2000US-00495406.
XX
XX (CAN ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;
XX
XX WPI; 2000-524418/47.
XX N-PSDB; AA53724.

```

```

XX
PT Novel glycosyltransferase polypeptides and polynucleotides useful for
PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
PT reagents and as immunogen for producing antibodies.
XX
PS Claim 13; Page 92-94; 120pp; English.
XX
CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
CC useful for synthesising sialylated oligosaccharides such as ganglioside,
CC lyso-ganglioside or their mimics. Glycosyltransferases are useful for
CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
CC other oligosaccharides that have biological activity. The enzymes and
CC nucleic acids that encode them are useful for studies of the pathogenesis
CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
CC jejuni and the nucleic acids are used as probes to study expression of
CC genes involved in ganglioside mimetic synthesis. Antibodies raised
CC against the glycosyltransferases are also useful for analysing the
CC expression patterns of these genes involved in pathogenesis. The nucleic
CC acids are also useful for designing antisense oligonucleotides for
CC inhibiting expression of the Campylobacter enzymes that are involved in
CC the biosynthesis of ganglioside mimics that can mask the pathogens from
CC the host's immune system. The oligosaccharides are useful as diagnosing
CC reagents or as therapeutics and as immunogens for producing antibodies.
CC Bacterial glycosyltransferase can be used to catalyse the formation of
CC oligosaccharides that are identical to the corresponding mammalian
CC structures and are easier and less expensive to produce in large
CC quantity, compared to the mammalian glycosyltransferase. The bacterial
CC origin of the enzymes facilitates expression of large quantities of the
CC enzymes using relatively inexpensive prokaryotic expression systems.
CC (Updated on 12-SEP-2003 to standardise OS field)
XX
SQ Sequence 291 AA;

Query Match 51.8%; Score 813; DB 3; Length 291;
Best Local Similarity 54.3%; Pred. No. 3.8e-66;
Matches 159; Conservative 46; Mismatches 70; Indels 18; Gaps 6;

QY 8 LVCNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 67
DB 5 ILAGNGPSLKEIDYSLRPNDVDFRCNQYFEDRYFVGKDVYFNFVFFFEQYYSKK 64
QY 68 LIQNEEYENIENIVCSTINLEYIDGQFVDNFELFSDAFILGHEIIKKLDFPFIYKNEI 127
DB 65 LIQNEEYETELIMCSNFNQAHLENQNFVKTIFYDFPDALHGYDFPKQKEFNAYFKFHEI 124
QY 128 YNQRITSGVYMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLLNKCTGFKNQKF 187
DB 125 YLNQRITSGVYMCATAVALGYSIYISGIDFYQNGSS-YAFDTKQENLLKLPDFKNDRS 183
QY 188 KFINHSMACDLQALDYLMKRYDVNIYSL--NS--DEYFKLAPDIGSDPVLSPKPKYIND 243
DB 184 HYIGHSKNTDIALKLEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK-NNYTKD 242
QY 244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLRLPSDIKHLYLKEK 288
DB 243 ILIPSSSEA----YGFKSNINFKIKIKENIYKLIKDLRLPSDIKHYPKKG 291

RESULT 8
ABJ18480
ID ABJ18480 standard; protein; 291 AA.
XX
AC ABJ18480;
XX
XX 07-FEB-2003 (first entry)
XX
DE Campylobacter jejuni bifunctional sialtransferase catII #2.
XX
XX Enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAC transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;

```

KW	ganglioside mimetics; inflammation; tumour metastasis.	DT	12-SEP-2003 (revised)	
XX		DT	22-DEC-2000 (first entry)	
XX	Campylobacter jejuni.	DE	Campylobacter jejuni O:19 serotype CstII sialyltransferase.	
XX	WO200274942-A2.	XX	Biosynthetic locus; biosynthesis; lipid A biosynthesis; acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase; Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase; sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody; immunity; immunogen; ganglioside.	
XX	26-SEP-2002.	OS	Campylobacter jejuni; O:19 serotype.	
XX	22-FEB-2002; 2002WO-CA000229.	XX	WO200046379-A1.	
XX	21-MAR-2001; 2001US-00816028.	PN	10-AUG-2000.	
XX	(CANA) NAT RES COUNCIL CANADA.	XX	01-FEB-2000; 2000WO-CA000086.	
XX	Gilbert M, Wakarchuk WW;	XX	01-FEB-1999; 99US-0118213P.	
XX	WPI; 2003-040554/03.	PR	31-JAN-2000; 2000US-00495406.	
DR	N-PSDB; ABT13667.	XX	(CANA) NAT RES COUNCIL CANADA.	
XX		PA	Gilbert M, Wakarchuk WW;	
XX	New glycosyltransferases from Campylobacter, useful for synthesizing gangliosides and ganglioside mimetics, and in studying the pathogenesis mechanisms of organisms that synthesize ganglioside mimetics.	PI	WPI; 2000-524418/47.	
XX	Claim 5; Page 97; 107pp; English.	XX	N-PSDB; AA53726.	
XX	The invention comprises the amino acid and coding sequences of Campylobacter jejuni proteins. The C. jejuni proteins of the invention may be either an: acyltransferase; glycosyltransferase; GalNAC (N-Acetyl-galactosamine) transferase; galactosyltransferase; sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present amino acid sequence represents a Campylobacter jejuni protein of the invention	PT	Novel glycosyltransferase polypeptides and polynucleotides useful for biosynthesis of ganglioside and ganglioside mimics, as diagnostic reagents and as immunogen for producing antibodies.	
XX		XX	Disclosure; Page 96-97; 120pp; English.	
XX	Sequence 291 AA;	CC	A reaction mixture for the synthesis of a sialylated oligosaccharide is useful for synthesizing sialylated oligosaccharides such as ganglioside, lyso-ganglioside or their mimics. Glycosyltransferases are useful for chemo-enzymatic synthesis of oligosaccharides, including gangliosides and other oligosaccharides that have biological activity. The enzymes and nucleic acids that encode them are useful for studies of the pathogenesis mechanisms of organisms that synthesize ganglioside mimics, such as C. jejuni and the nucleic acids are used as probes to study expression of genes involved in ganglioside mimetic synthesis. Antibodies raised against the glycosyltransferases are also useful for analyzing the expression patterns of these genes involved in pathogenesis. The nucleic acids are also useful for designing antisense oligonucleotides for inhibiting expression of the Campylobacter enzymes that are involved in the biosynthesis of ganglioside mimics that can mask the pathogens from the host's immune system. The oligosaccharides are useful as diagnosing reagents or as therapeutics and as immunogens for producing antibodies. Bacterial glycosyltransferase can be used to catalyse the formation of oligosaccharides that are identical to the corresponding mammalian structures and are easier and less expensive to produce in large quantity, compared to the mammalian glycosyltransferase. The bacterial origin of the enzymes facilitates expression of large quantities of the enzymes using relatively inexpensive prokaryotic expression systems. (Updated on 12-SEP-2003 to standardise OS field)	
XX		XX	Sequence 291 AA;	
XX		XX	Query Match 51.2%; Score 804; DB 3; Length 291;	
XX		XX	Best Local Similarity 53.9%; Pred. No. 2.6e-65;	
XX		XX	Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;	
QY	8 LVCNGSPSLKNIIDYKRLPKQDFVRCNQFYEDRYFGKDVKYVFNPFVFFQYTSKK 67	QY	8 LVCNGSPSLKNIIDYKRLPKQDFVRCNQFYEDRYFGKDVKYVFNPFVFFQYTSKK 67	
Db	5 IIAAGNPSLKEIDYSRLPNDVFRNCQFYEDKYYLGGKCAVFNPGGLFFEQYTLKH 64	Db	5 IIAAGNPSLKEIDYSRLPNDVFRNCQFYEDKYYLGGKCAVFNPGGLFFEQYTLKH 64	
QY	68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDFFAYIKYNEI 127	QY	68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDFFAYIKYNEI 127	
Db	65 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDFFAYIKYNEI 124	Db	65 LIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDFFAYIKYNEI 124	
QY	128 YNRQRTSGVYMCATAVALGYKSIYSGIDFYQDTNNLYAFDNNKNLKNCTGFRNQK 187	QY	128 YNRQRTSGVYMCATAVALGYKSIYSGIDFYQDTNNLYAFDNNKNLKNCTGFRNQK 187	
Db	125 YLNQRTSGVYMCATAVALGYKSIYSGIDFYQNGSS-YAFDTKQENLLKLAPDFKND 183	Db	125 YLNQRTSGVYMCATAVALGYKSIYSGIDFYQNGSS-YAFDTKQENLLKLAPDFKND 183	
QY	188 KFINHSMACDLQALDYLMKDYDNIYSL--NS--DEYFKLAPDIGDFVLSPKPKYIND 243	QY	188 KFINHSMACDLQALDYLMKDYDNIYSL--NS--DEYFKLAPDIGDFVLSPKPKYIND 243	
Db	184 HYIGHSKNTDIKALEFLKTYKIKLYCLPCNSLANFIELAPLNSNFIIQEK--NNYTKD 242	Db	184 HYIGHSKNTDIKALEFLKTYKIKLYCLPCNSLANFIELAPLNSNFIIQEK--NNYTKD 242	
QY	244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLRLPSDKHYLKEK 288	QY	244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLRLPSDKHYLKEK 288	
Db	243 ILIPSESA----YKTSKNINFKKIKENIYYKLIKDLRLPSDKHYLKEK 291	Db	243 ILIPSESA----YKTSKNINFKKIKENIYYKLIKDLRLPSDKHYLKEK 291	
RESULT 9		RESULT 9		
AA97212		AA97212		
ID	AA97212 standard; protein; 291 AA.	ID	AA97212 standard; protein; 291 AA.	
XX		XX		
AC	AA97212;	AC	AA97212;	
XX		XX		

QY 128 YNRQRTSGVTMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLKNKCTGFKNQKF 187
 DB 125 YFNQRTSGVTMCATAVALGYSIYISGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183
 QY 188 KFINHSMACDQALDYLKRYDNIYSL--NS--DEYFKLAPDIGDFVLSSKPKKYIND 243
 DB 184 HYIGHSKNTDIKALEPLEKTKYKLYCLCPNSLLANFIELAPLNSNFIQEK-NNYTKD 242
 QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLIKDLIRLPSDIKHLYKEK 288
 DB 243 ILIPSSSEA----YGFKSKNINFKIKIKENYVYKLIKDLIRLPSDIKHYPKKG 291

RESULT 10
 ABJ18485
 ID ABJ18485 standard; protein; 291 AA.
 XX
 AC ABJ18485;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Campylobacter jejuni bifunctional sialtransferase catII #6.
 XX
 KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.
 XX
 OS Campylobacter jejuni.
 XX
 PN WO200274942-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-CA000229.
 XX
 PR 21-MAR-2001; 2001US-00816028.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Gilbert M, Wakarchuk WW;
 XX
 DR WPI; 2003-040554/03.
 DR N-PSDB; ABT13671.
 XX
 PT New glycosyltransferases from Campylobacter, useful for synthesizing
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 PT mechanisms of organisms that synthesize ganglioside mimetics.
 XX
 PS Disclosure; Page 99; 107pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
 CC Acetylgalactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of Campylobacter enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. jejuni oligosaccharides of the invention may be used as
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
 CC metastasis). The present amino acid sequence represents a Campylobacter
 CC jejuni protein of the invention

Sequence 291 AA;
 Query Match 51.2%; Score 804; DB 6; Length 291;
 Best Local Similarity 53.9%; Pred. No. 2.6e-65;
 Matches 158; Conservative 45; Mismatches 72; Indels 18; Gaps 6;

QY 8 LVCNGPSPKNDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFPNPFVFPQYYSKK 67
 DB 5 IIAAGPSLKEIDYSRLPNDVFRCNQFYFEDRYFVGKDVKYVFPNPFVFPQYYSKK 64
 QY 68 LIQNEBYNTENIVCSNTINLEYIDGPOFVDFNFELYSFSDAFGLGHEIIKKLQDFRAYIKYNEI 127
 DB 65 LIQNEBYNTENIVCSNTINLEYIDGPOFVDFNFELYSFSDAFGLGHEIIKKLQDFRAYIKYNEI 124
 QY 128 YNRQRTSGVTMCATAVALGYSIYISGIDFYQDTNNLYAFDNNKNLKNKCTGFKNQKF 187
 DB 125 YFNQRTSGVTMCATAVALGYSIYISGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183
 QY 188 KFINHSMACDQALDYLKRYDNIYSL--NS--DEYFKLAPDIGDFVLSSKPKKYIND 243
 DB 184 HYIGHSKNTDIKALEPLEKTKYKLYCLCPNSLLANFIELAPLNSNFIQEK-NNYTKD 242
 QY 244 ILIPDKYAQERYVG-----KKSRLKENLHYKLIKDLIRLPSDIKHLYKEK 288
 DB 243 ILIPSSSEA----YGFKSKNINFKIKIKENYVYKLIKDLIRLPSDIKHYPKKG 291

RESULT 11
 ABJ18484
 ID ABJ18484 standard; protein; 291 AA.
 XX
 AC ABJ18484;
 XX
 DT 07-FEB-2003 (first entry)
 XX
 DE Campylobacter jejuni bifunctional sialtransferase catII #5.
 XX
 KW Enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylgalactosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.
 XX
 OS Campylobacter jejuni.
 XX
 PN WO200274942-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 22-FEB-2002; 2002WO-CA000229.
 XX
 PR 21-MAR-2001; 2001US-00816028.
 XX
 PA (CANADA) NAT RES COUNCIL CANADA.
 XX
 PI Gilbert M, Wakarchuk WW;
 XX
 DR WPI; 2003-040554/03.
 DR N-PSDB; ABT13670.
 XX
 PT New glycosyltransferases from Campylobacter, useful for synthesizing
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 PT mechanisms of organisms that synthesize ganglioside mimetics.
 XX
 PS Disclosure; Page 99; 107pp; English.
 XX
 CC The invention comprises the amino acid and coding sequences of
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
 CC Acetylgalactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of Campylobacter enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. jejuni oligosaccharides of the invention may be used as

PS Claim 27; Fig 2; 47pp; English.
 XX
 CC The present sequence represents Campylobacter jejuni alpha-2,3-sialyltransferase which is encoded by the cst-I gene. The alpha-2,3-sialyltransferase protein is useful for producing desired carbohydrate structures by contacting the acceptor molecule (which has a terminal galactose residue) with an activated sialic acid molecule. The terminal galactose residue is linked to a second residue (Glc or a GlcNac, or GlcNac or GalNac) in the acceptor molecule through a beta-1,3 or beta-1,4 linkage, respectively. The activated sialic acid is CMP-Neu5Ac. The CC polynucleotides and polypeptides facilitate the improved production of CC desired structures and nucleic acids encoding sialyltransferases
 XX
 SQ Sequence 430 AA;
 Query Match 43.3%; Score 680.5; DB 2; Length 430;
 Best Local Similarity 49.7%; Pred. No. 9.1e-54;
 Matches 145; Conservative 49; Mismatches 75; Indels 23; Gaps 8;
 QY 1 MSMMI-NALVCGNGPSLKNIDYKRLPKQDFVFCRCNQFYFEDRYVGVKDYVFNPFVFF 59
 DB 12 VSKMQMIIAGNGPSLKNIDYKRLPKQDFVFCRCNQFYFEDRYVGVKDYVFNPFVFF 71
 QY 60 EGYTSKLLIQNEEYNIENIVCSTINLEYIDGQFVDNFEELYSDFALGHEIILKLDFF 119
 DB 72 QVHTAKQLIKNIEYEIKNIFCSFNPFIESNDFLHQFYNFPDPAKLGVEVIENLKEFY 131
 QY 120 AYIKNEIYNRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNK 179
 DB 132 AYIKNEIYNKRITSGVYMCATAVALGYKSIYILCGIDFYEG-DVIYFFAMSTNIKIF 190
 QY 180 TGFKNQKFKTIN-HSMACDQALDYLKMKRYDVNIYSLNSD-----EYFKLAPDIGSDPVL 234
 DB 191 PGIND--FKPSNCHSKSEYDEALKLSIYKVNIALCDDSIILANHPPLSININNNFTLE 248
 QY 235 KPKKYINDILIPDKYAQERYGKSRLEN-----LHYK--LIKDL 274
 DB 249 NKNNSINDILLTNTPTGVSFY--KNQKADKNKIMLNFYILHSKONLIKPL 298
 RESULT 14
 ABR40195
 ID ABR40195 standard; protein; 303 AA.
 AC ABR40195;
 XX
 XX 23-JUL-2003 (first entry)
 DT alpha-2,3/alpha-2,8-sialyltransferase #3.
 DE alpha-2,3/alpha-2,8-sialyltransferase #3.
 KW alpha-2,3/alpha-2,8-sialyltransferase; enzyme; sialic acid.
 XX Pasteurella multocida.
 OS
 XX WO2003027297-A1.
 PN
 XX 03-APR-2003.
 PD
 XX 26-SEP-2002; 2002WO-JP009907.
 PP
 XX 26-SEP-2001; 2001JP-00292796.
 PR
 XX (KYOW) KYOWA HAKKO KOGYO KK.
 PA
 XX Endo T, Koizumi S;
 PI
 XX WPI; 2003-393339/37.
 DR N-PSDB; ACC71694.
 XX
 XX Alpha 2,3/alpha 2,8 sialyltransferase from Pasteurella multocida for
 PT production of sialic acid-containing complex sugars.
 PT
 XX

PS Claim 9; Page 60-62; 66pp; Japanese.
 XX
 CC The present invention relates to a method for producing alpha-2,3/alpha-2,8-sialyltransferase. alpha-2,3/alpha-2,8-sialyltransferase can be used in a method for producing sialic acid-containing complex sugars from an oligosaccharide substrate using a culture of a microorganism transformed with DNA encoding alpha-2,3/alpha-2,8-sialyltransferase, or an extract of the culture. The present sequence is a protein sequence for alpha-2,3/alpha-2,8-sialyltransferase from Pasteurella multocida
 XX
 SQ Sequence 303 AA;
 Query Match 42.9%; Score 673.5; DB 6; Length 303;
 Best Local Similarity 46.8%; Pred. No. 2.6e-53;
 Matches 137; Conservative 47; Mismatches 94; Indels 15; Gaps 7;
 QY 8 LVCGNGPSLKNIDYKRLPKQDFVFCRCNQFYFEDRYVGVKDYVFNPFVFFEQYTSKK 67
 DB 14 IVAGNGESLSQIDYRLLPKNYDVFCRCNQFYFERYFLGNKIKAVFTPGVFLQYTYLYH 73
 QY 68 LIQNEEYNIENIVCSTINLEYIDGQFVDNFEELYSDFALGHE-IKKLKDFFAYIKNE 126
 DB 74 LKRNNEFYVDNVILSSFNHPTVD-LEKSQKIQALFIDVINGYEKLSKLTAFDVLRYKE 132
 QY 127 IYNRQITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQK 186
 DB 133 LYENQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNKCTGFKNQK 192
 QY 187 FKFINHSMACDQALDYLKMKRYDVNIYSLNS-----DYEYFKLAPDIGSD----FVLSKKPK 238
 DB 193 TLESYHSKIDLEALSFLQOHYHVNYSISPMSPSKHPPI-PTVEDDCETTFVAPLK-E 250
 QY 239 KYINDILIPDKYAQER---YYGKSKLENLHYKLIKDLIRLPSDIKHYLKEK 288
 DB 251 NYINDILLPPHFVYKLGITVSKSRFHSNLIVRLIRDLLKPSALKHYLKEK 303
 RESULT 15
 AAB18265
 ID AAB18265 standard; protein; 2013 AA.
 XX
 AC AAB18265;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 XX Plasmodium falciparum chromosome 2 related protein SEQ ID NO:122.
 DE
 KW Plasmodium falciparum; chromosome 2; human malaria parasite; vaccine; antimalarial; malaria; protozoicide; infection; insecticide.
 XX
 OS Plasmodium falciparum.
 XX WO200025728-A2.
 PN
 XX 11-MAY-2000.
 PD
 XX 05-NOV-1999; 99WO-US026796.
 PP
 XX 05-NOV-1998; 98US-0107131P.
 PR
 XX (HOFF/) HOFFMAN S.
 PA (CARU/) CARUCCI D.
 PA (GARD/) GARDNER M.
 PA (VENT/) VENTER J C.
 XX
 XX Hoffman S, Carucci D, Gardner M, Venter JC;
 PI
 XX WPI; 2000-365347/31.
 DR
 XX
 XX Proteins encoded by chromosome 2 of the human malarial parasite, PT Plasmodium falciparum, useful as antimalarial vaccines and in the PT diagnosis of P.falciparum infection.
 PT
 XX

PS Disclosure; Page 285-291; 577pp; English.

XX

The present invention describes proteins and their fragments (I) encoded by chromosome 2 of the human malarial parasite, Plasmodium falciparum. Also described are: (I) nucleotide sequences (II) encoding (I); and (2) vaccines against P. falciparum infection comprising (I) or (II). (I) and (II) are useful for the development of vaccines against P. falciparum infection. (I) and polyclonal antisera or a monoclonal antibody raised to immunogens comprising the sequences of (I), are useful in the detection of infection with P. falciparum. Furthermore, (I) are useful when they are rifins or secreted or membrane proteins) can aid the identification of drugs to treat or prevent P. falciparum infection, or they can be used to identify drug resistance in P. falciparum. Sequencing of the Plasmodium chromosome 2 and the subsequent identification of proteins encoded by it will help to expand our understanding of parasite biology, a process hampered by the complexity of the parasitic lifecycle, and provide new targets for vaccine and drug development. Parasite resistance to drugs and mosquito resistance to insecticides have led to a resurgence of malaria in many parts of the world, and there is a pressing need for vaccines and new drugs. AAA70078 to AAA70287 and AAB18144 to AAB18352 represent nucleotide and protein sequences given in the present invention, but which are not specifically mentioned within the specification

XX
SQ

Sequence 2013 AA;

Query Match		8.2%	Score 128.5;	DB 3;	Length 2013;
Best Local Similarity		18.9%	Pred. No. 0.03;		
Matches		66;	Conservative	59;	Mismatches 98; Indels 127; Gaps 15;
Qy	34	NQFYFEDRYVGVKDYKVFNFVFFFE--QYVTSKKLIQNEEYNIEN-----IVCSTINL	86		
Db	1147	NKTEFNDNMIKEDKLEEKINEDFVITEGEKSKNKKIKNTQHNNDNNNDVFCNSL--	1204		
Qy	87	EYIDGQFQVDFELYFSDAFLGHEIITKLDKDPFAYIKYNEI--YRQRI-----	133		
Db	1205	-----YEL-----LLNKEKSFFLNIKHGKLYINERMHTSELTYYIDIVT	1243		
Qy	134	TSGVVMCATAVAGYK-----SIVISGIDFYQDTNNLY--APD	169		
Db	1244	TNNILICISFNSVDYPLEINPHINIRMPYLNNDIQYYPLIILKGNNNNNMYDLFL	1303		
Qy	170	NNKKNLLNKTGFRNQ-----KPKFINHSMACDLQALD-----YLMKRYDVIY--SL	215		
Db	1304	IKKKNFLLLRNNIKEDBEAIIKQEKDHSITCNPKLIQNCNDQTYNTKCVSENVFNTI	1363		
Qy	216	NSDE-----YFKLAPDIGSDFVLSKPKKYINDI	244		
Db	1364	NSNEHISFYLSKWIIEDNNTSYINDSLIKNMNIVFLKIKNDISQNYT--NRKEKNFFDI	1422		
Qy	245	LIPDKVAQERYYGKSKRLKENL-----HYKLIKDLIRLPSPDIK	282		
Db	1423	VCKEKKYIEN--NKNNEKNKIVDINNNMPTHYNIILKNKILLNDVE	1470		

Search completed: June 13, 2005, 21:06:23
Job time : 164 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 21:00:06 ; Search time 43 Seconds
(without alignments)
510.391 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSNNINALVCGNGPSLKNIID.....IRLPDIKHYLKEKYANKNR 294

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	294	4	US-09-495-406-10
2	1571	100.0	294	4	US-09-816-028A-10
3	1571	100.0	294	4	US-10-303-162-10
4	1571	100.0	294	4	US-10-303-134-10
5	823	52.4	291	4	US-09-495-406-3
6	823	52.4	291	4	US-09-816-028A-3
7	823	52.4	291	4	US-10-303-162-3
8	823	52.4	291	4	US-10-303-134-3
9	817	52.0	291	4	US-09-495-406-7
10	817	52.0	291	4	US-09-816-028A-7
11	817	52.0	291	4	US-10-303-162-7
12	817	52.0	291	4	US-10-303-134-7
13	813	51.8	291	4	US-09-495-406-5
14	813	51.8	291	4	US-09-816-028A-5
15	813	51.8	291	4	US-10-303-162-5
16	813	51.8	291	4	US-10-303-134-5
17	804	51.2	291	4	US-09-495-406-9
18	804	51.2	291	4	US-09-816-028A-9
19	804	51.2	291	4	US-09-816-028A-12
20	804	51.2	291	4	US-09-816-028A-14
21	804	51.2	291	4	US-10-303-162-9
22	804	51.2	291	4	US-10-303-162-12
23	804	51.2	291	4	US-10-303-162-14
24	804	51.2	291	4	US-10-303-134-9
25	804	51.2	291	4	US-10-303-134-12
26	804	51.2	291	4	US-10-303-134-14
27	680.5	43.3	322	4	US-09-495-406-34

28 680.5 43.3 322 4 US-09-816-028A-48 Sequence 48, Appl
29 680.5 43.3 322 4 US-10-303-162-48 Sequence 48, Appl
30 680.5 43.3 322 4 US-10-303-134-48 Sequence 48, Appl
31 680.5 43.3 430 4 US-09-272-960-2 Sequence 2, Appl
32 680.5 43.3 430 4 US-10-058-636-2 Sequence 2, Appl
33 441.5 28.1 231 4 US-09-495-406-35 Sequence 35, Appl
34 441.5 28.1 231 4 US-09-272-960-5 Sequence 5, Appl
35 441.5 28.1 231 4 US-09-816-028A-49 Sequence 49, Appl
36 441.5 28.1 231 4 US-10-058-636-5 Sequence 5, Appl
37 441.5 28.1 231 4 US-10-303-162-49 Sequence 49, Appl
38 441.5 28.1 231 4 US-10-303-134-49 Sequence 49, Appl
39 112 7.1 1169 4 US-09-583-110-4409 Sequence 4409, Ap
40 112 7.1 1169 4 US-09-107-433-4847 Sequence 4847, Ap
41 104 6.6 367 4 US-09-861-451A-46 Sequence 46, Appl
42 104 6.6 898 1 US-08-465-995A-4 Sequence 4, Appl
43 104 6.6 898 2 US-08-465-994C-4 Sequence 4, Appl
44 104 6.6 898 2 US-08-966-145-4 Sequence 4, Appl
45 104 6.6 920 1 US-08-101-593-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-495-406-10
; Sequence 10, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ IDS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-09-495-406-10

Query Match 100.0%; Score 1571; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-153;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSNNINALVCGNGPSLKNIIDYKELPKQFVFCRNCQFVEDRYFVGKDVKYVFENPFVFFE 60
Db 1 MSNNINALVCGNGPSLKNIIDYKELPKQFVFCRNCQFVEDRYFVGKDVKYVFENPFVFFE 60
Qy 61 QYYSKKLIQNEBYNIENIVCSTINLEYIDGQFVDNMFELYFSDAFLGHEIHKLDKFFA 120
Db 61 QYYSKKLIQNEBYNIENIVCSTINLEYIDGQFVDNMFELYFSDAFLGHEIHKLDKFFA 120
Qy 121 YIKYNIYNQRITSGVYMCATAVALGYKSIYISIGIDFYODTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNIYNQRITSGVYMCATAVALGYKSIYISIGIDFYODTNNLYAFDNNKKNLNKCT 180
Qy 181 GFKNQKFKETINHSMACDLQALDYLMKRYDVNLYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db 181 GFKNQKFKETINHSMACDLQALDYLMKRYDVNLYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy 241 INDILIPDKYAQERYYGKESRLKENLHYKLIKDILRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYYGKESRLKENLHYKLIKDILRLPSDIKHYLKEKYANKNR 294

```
RESULT 2
US-09-816-028A-10
; Sequence 10, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-09-816-028A-10
Query Match 100.0%; Score 1571; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-153;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVYFNFVFPFVE 60
Db 1 MSMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVYFNFVFPFVE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHEIIKKLDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHEIIKKLDFFA 120
Qy 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Db 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYANKNR 294
RESULT 3
US-10-303-162-10
; Sequence 10, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-162-10
Query Match 100.0%; Score 1571; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-153;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVYFNFVFPFVE 60
Db 1 MSMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVYFNFVFPFVE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHEIIKKLDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNPFYSDAFLGHEIIKKLDFFA 120
Qy 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Db 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKYANKNR 294
RESULT 4
US-10-303-134-10
; Sequence 10, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-134-10
Query Match 100.0%; Score 1571; DB 4; Length 294;
Best Local Similarity 100.0%; Pred. No. 1e-153;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MSMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVYFNFVFPFVE 60
Db 1 MSMNALVCGNGPSLKNIDYKRLPKQFDVFRNCQYFEDRYFVGKDVYFNFVFPFVE 60
```

QY 61 QYTSKLIQNEEYNIIVCSTINLEYIDGQFQVDFNFELYSDFDAFLGHIHIIKLDKFFA 120
 Db 61 QYTSKLIQNEEYNIIVCSTINLEYIDGQFQVDFNFELYSDFDAFLGHIHIIKLDKFFA 120
 QY 121 YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 Db 121 YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCT 180
 QY 181 GFKNQKFKFNHSMACDLQALDYLKRYDNIIVSLNSDEYFKLAPDYGSDVLSKPKKY 240
 Db 181 GFKNQKFKFNHSMACDLQALDYLKRYDNIIVSLNSDEYFKLAPDYGSDVLSKPKKY 240
 QY 241 INDILIPDKYAOERYGKSKRLKLNHLYKLIKDLIRLPSDIKHLYKEKYANKR 294
 Db 241 INDILIPDKYAOERYGKSKRLKLNHLYKLIKDLIRLPSDIKHLYKEKYANKR 294

RESULT 5

US-09-495-406-3

; Sequence 3, Application US/09495406
 ; Patent No. 6503744

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; FILE OF INVENTION: Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-000110US

; CURRENT APPLICATION NUMBER: US/09/495,406

; CURRENT FILING DATE: 2000-01-31

; PRIOR APPLICATION NUMBER: US 60/118,213

; PRIOR FILING DATE: 1999-02-01

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 291

; TYPE: PRT

; ORGANISM: Campylobacter jejuni

; US-09-495-406-3

Query Match 52.4%; Score 823; DB 4; Length 291;
 Best Local Similarity 54.9%; Pred. No. 1.4e-76;
 Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFFSQYTSKK 67
 Db 5 IIAAGNGPSLKEIDYSLPNDPFDVRCNQFYFEDRYFVGKDVKYVFNPFVFFSQYTSKK 64
 QY 68 LIQNEEYNIIVCSTINLEYIDGQFQVDFNFELYSDFDAFLGHIHIIKLDKFFAIIKNEI 127
 Db 65 LIQNEEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFPKQLKDFNAYPKFHEI 124
 QY 128 YNRQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCTGFKNQKF 187
 Db 125 YFNQRITSGVYMCATAVALGYKSIYISGIDFYQNGSS-YAFDTKQKLLKLANFNFDNS 183
 QY 188 KFINHSMACDLQALDYLKRYDNIIVSL--NS--DEYFKLAPDYGSDVLSKPKKYIND 243
 Db 184 HYIGHSKNTDIKALEPLEKTYKIKLYCLPCNSLLANFIELAPLNSNFIQEK--NNYTKD 242
 QY 244 ILIPDKYAOERYG-----KKSRLKENLHLYKLIKDLIRLPSDIKHLYKEK 288
 Db 243 ILIPSSA----YGRFSKNINFKIKIKENIYKLIKDLIRLPSDIKHLYKEK 291

RESULT 6

US-09-816-028A-3

; Sequence 3, Application US/09816028A

; Patent No. 6693705

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/09/816,028A
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 291
 ; TYPE: PRT
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 ; OTHER INFORMATION: biosynthesis locus)
 ; US-09-816-028A-3

Query Match 52.4%; Score 823; DB 4; Length 291;

Best Local Similarity 54.9%; Pred. No. 1.4e-76;

Matches 161; Conservative 43; Mismatches 71; Indels 18; Gaps 6;

QY 8 LVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVKYVFNPFVFFSQYTSKK 67
 Db 5 IIAAGNGPSLKEIDYSLPNDPFDVRCNQFYFEDRYFVGKDVKYVFNPFVFFSQYTSKK 64
 QY 68 LIQNEEYNIIVCSTINLEYIDGQFQVDFNFELYSDFDAFLGHIHIIKLDKFFAIIKNEI 127
 Db 65 LIQNEEYETELIMCSNNOAHLENENFVKTFFDYFDPDAHLGYDFPKQLKDFNAYPKFHEI 124
 QY 128 YNRQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCTGFKNQKF 187
 Db 125 YFNQRITSGVYMCATAVALGYKSIYISGIDFYQNGSS-YAFDTKQKLLKLANFNFDNS 183
 QY 188 KFINHSMACDLQALDYLKRYDNIIVSL--NS--DEYFKLAPDYGSDVLSKPKKYIND 243
 Db 184 HYIGHSKNTDIKALEPLEKTYKIKLYCLPCNSLLANFIELAPLNSNFIQEK--NNYTKD 242
 QY 244 ILIPDKYAOERYG-----KKSRLKENLHLYKLIKDLIRLPSDIKHLYKEK 288
 Db 243 ILIPSSA----YGRFSKNINFKIKIKENIYKLIKDLIRLPSDIKHLYKEK 291

RESULT 7

US-10-303-162-3

; Sequence 3, Application US/10303162

; Patent No. 6723545

; GENERAL INFORMATION:

; APPLICANT: Gilbert, Michel

; APPLICANT: Wakarchuk, Warren W.

; APPLICANT: National Research Council of Canada

; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; FILE OF INVENTION: Gangliosides and Ganglioside Mimics

; FILE REFERENCE: 019633-000111US

; CURRENT APPLICATION NUMBER: US/10/303,162

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US/09/816,028

; PRIOR FILING DATE: 2001-03-21

; PRIOR APPLICATION NUMBER: US 60/118,213

; PRIOR FILING DATE: 1999-02-01

; PRIOR APPLICATION NUMBER: US 09/495,406

; NUMBER OF SEQ ID NOS: 49

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 3

; LENGTH: 291

; TYPE: PRT

; ORGANISM: Campylobacter jejuni


```
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-09-816-028A-7

Query Match      52.0%; Score 817; DB 4; Length 291;
Best Local Similarity 54.3%; Pred. No. 5.7e-76;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY      8 LVCNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGDKVYFFNPFFVFEQYTSKK 67
Db      5 ILAGNGPSLKEIDYSLRPNDQDFVRCNQYFEDRYFVGDKVYFFNPFFVFEQYTSKK 64
QY      68 LIQNEEYNIENIVCSTINLEYIDGQFQVDNPFYFSDAFGLGHEIIKKLDKDFPAYIKYNEI 127
Db      65 LIQNEEYETELIMCSNFNOAHLENQNFVTFDYDPDAHLGYDFFKQLKEFNAYFKFHEI 124
QY      128 YNRQRTISGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFRNQKF 187
Db      125 YFNQRTISGVYMCATAVALGYKSIYISGIDFYQDNGSS-YAFDTKQKLLKLAFLNFKDNS 183
QY      188 KFINHSMACDLQALDYLKRYDVIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
Db      184 HYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
QY      244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
Db      243 ILIPSSA-----YGFKTKNINFKIKENIYKLIKDLIRLPSDIKHYFKGK 291

RESULT 11
US-10-303-162-7
; Sequence 7, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-303-162-7

Query Match      52.0%; Score 817; DB 4; Length 291;
Best Local Similarity 54.3%; Pred. No. 5.7e-76;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY      8 LVCNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGDKVYFFNPFFVFEQYTSKK 67
Db      5 ILAGNGPSLKEIDYSLRPNDQDFVRCNQYFEDRYFVGDKVYFFNPFFVFEQYTSKK 64
QY      68 LIQNEEYNIENIVCSTINLEYIDGQFQVDNPFYFSDAFGLGHEIIKKLDKDFPAYIKYNEI 127
Db      65 LIQNEEYETELIMCSNFNOAHLENQNFVTFDYDPDAHLGYDFFKQLKEFNAYFKFHEI 124
QY      128 YNRQRTISGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFRNQKF 187
Db      125 YFNQRTISGVYMCATAVALGYKSIYISGIDFYQDNGSS-YAFDTKQKLLKLAFLNFKDNS 183
QY      188 KFINHSMACDLQALDYLKRYDVIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
Db      184 HYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
QY      244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
Db      243 ILIPSSA-----YGFKTKNINFKIKENIYKLIKDLIRLPSDIKHYFKGK 291

RESULT 12
US-10-303-134-7
; Sequence 7, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 7
; LENGTH: 291
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-10-303-134-7

Query Match      52.0%; Score 817; DB 4; Length 291;
Best Local Similarity 54.3%; Pred. No. 5.7e-76;
Matches 159; Conservative 45; Mismatches 71; Indels 18; Gaps 6;

QY      8 LVCNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGDKVYFFNPFFVFEQYTSKK 67
Db      5 ILAGNGPSLKEIDYSLRPNDQDFVRCNQYFEDRYFVGDKVYFFNPFFVFEQYTSKK 64
QY      68 LIQNEEYNIENIVCSTINLEYIDGQFQVDNPFYFSDAFGLGHEIIKKLDKDFPAYIKYNEI 127
Db      65 LIQNEEYETELIMCSNFNOAHLENQNFVTFDYDPDAHLGYDFFKQLKEFNAYFKFHEI 124
QY      128 YNRQRTISGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKQNLNCKCTGFRNQKF 187
Db      125 YFNQRTISGVYMCATAVALGYKSIYISGIDFYQDNGSS-YAFDTKQKLLKLAFLNFKDNS 183
QY      188 KFINHSMACDLQALDYLKRYDVIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
Db      184 HYIGHSKNTDIKALEFLEKTYEIKLYCLCPNSLLANFIELAPNLSNFIQEK--NNYTKD 242
QY      244 ILIPDKYAQERYYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
Db      243 ILIPSSA-----YGFKTKNINFKIKENIYKLIKDLIRLPSDIKHYFKGK 291
```



```

Qy 8 LVCNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVKDKVKYVFNPFVFFQYITSKK 67
Db : : ||||| ||||| ||||| ||||| : : || : || : ||||| |
5 ILAGNPSLKEIDYSRLENDFVRCNQYFEDKYLGKKFKAVFYNPGLFFQYITLKH 64
Qy 68 LIQNEYNINIENIVCSTINLEYIDGFQVDFNPFYSDAFLGHEIIKKLKDFFAYIKYNEI 127
Db ||||| ||||| : : || : || : ||||| ||||| : : || : || : ||||| |
65 LIQNEYETELIMCSNYNQAHLENENFVKTFFDPAHLGYDFFKQKFNAYFKFHEI 124
Qy 128 YNRORITSGVVMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLLNKCTGFRNQKF 187
Db | ||||| ||||| : ||||| : ||||| : : ||||| : : ||||| : ||||| : ||||| :
125 YLNQRITSGVVMCAVALGYKEIYLSGIDFYQNGSS-YAFDTKQENLLKLAPDFKNDRS 183
Qy 188 KFINHSMACDLQALDYLKRYDVNIYSL--NS--DEYFKLAPDIGSDVLSKKPKKYIND 243
Db : || : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
184 HYIGHSKNTDIKALEFLEKTYKIKYCLCPNSLLANFIELAPNLSNFIIOEK-NNYTKD 242
Qy 244 ILIPDKYAQERYG-----KKSRLKENLHYKLIKDLIRLPSDIKHYLKEK 288
Db ||||| ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
243 ILIPSSA-----YOKPSKNINPKIKIKENIYYKLIKDLIRLPSDIKHYFKGK 291

```

Search completed: June 13, 2005, 21:10:56
Job time : 44 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2005, 21:09:32 ; Search time 157 Seconds
(without alignments)
717.836 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNALVCGGSLKNID.....IRLPDIKHVKEKANKR 294

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1571	100.0	294	9	US-09-816-028A-10
2	1571	100.0	294	14	US-10-303-161-10
3	1571	100.0	294	14	US-10-303-118-10
4	1571	100.0	294	14	US-10-303-128-10
5	1571	100.0	294	14	US-10-303-134-10
6	1571	100.0	294	14	US-10-303-162-10
7	1571	100.0	294	16	US-10-845-408-10
8	1571	100.0	294	16	US-10-845-408-10
9	1571	100.0	294	16	US-10-845-412-10
10	1571	100.0	294	16	US-10-846-219-10
11	1571	100.0	294	16	US-10-821-604-10

12	1571	100.0	294	16	US-10-847-983-10	Sequence 10, Appl
13	1571	100.0	294	16	US-10-821-573-10	Sequence 10, Appl
14	1571	100.0	294	16	US-10-850-807-10	Sequence 10, Appl
15	1571	100.0	294	16	US-10-850-125-10	Sequence 10, Appl
16	1571	100.0	294	16	US-10-830-825-10	Sequence 10, Appl
17	1571	100.0	294	17	US-10-962-334-10	Sequence 10, Appl
18	1571	100.0	294	17	US-10-830-997-10	Sequence 10, Appl
19	1571	100.0	294	17	US-10-962-235-10	Sequence 10, Appl
20	823	52.4	291	9	US-09-816-028A-3	Sequence 3, Appl
21	823	52.4	291	14	US-10-303-161-3	Sequence 3, Appl
22	823	52.4	291	14	US-10-303-118-3	Sequence 3, Appl
23	823	52.4	291	14	US-10-303-128-3	Sequence 3, Appl
24	823	52.4	291	14	US-10-303-134-3	Sequence 3, Appl
25	823	52.4	291	14	US-10-303-162-3	Sequence 3, Appl
26	823	52.4	291	16	US-10-820-536-3	Sequence 3, Appl
27	823	52.4	291	16	US-10-845-408-3	Sequence 3, Appl
28	823	52.4	291	16	US-10-845-412-3	Sequence 3, Appl
29	823	52.4	291	16	US-10-846-219-3	Sequence 3, Appl
30	823	52.4	291	16	US-10-821-604-3	Sequence 3, Appl
31	823	52.4	291	16	US-10-847-983-3	Sequence 3, Appl
32	823	52.4	291	16	US-10-821-573-3	Sequence 3, Appl
33	823	52.4	291	16	US-10-850-807-3	Sequence 3, Appl
34	823	52.4	291	16	US-10-850-125-3	Sequence 3, Appl
35	823	52.4	291	16	US-10-830-825-3	Sequence 3, Appl
36	823	52.4	291	17	US-10-962-334-3	Sequence 3, Appl
37	823	52.4	291	17	US-10-830-997-3	Sequence 3, Appl
38	823	52.4	291	17	US-10-962-235-3	Sequence 3, Appl
39	817	52.0	291	9	US-09-816-028A-7	Sequence 7, Appl
40	817	52.0	291	14	US-10-303-161-7	Sequence 7, Appl
41	817	52.0	291	14	US-10-303-118-7	Sequence 7, Appl
42	817	52.0	291	14	US-10-303-128-7	Sequence 7, Appl
43	817	52.0	291	14	US-10-303-134-7	Sequence 7, Appl
44	817	52.0	291	14	US-10-303-162-7	Sequence 7, Appl
45	817	52.0	291	16	US-10-820-536-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-816-028A-10
; Sequence 10, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816, 028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495, 406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-09-816-028A-10

Query Match 100.0%; Score 1571; DB 9; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSNNALVCGGSLKNIDYKELPKQFDVFCNQFVFKDKVKKVFFVFFVFF 60
|||||

Db 1 MSNNALVCGNPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIILKLDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIILKLDFFA 120
Qy 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Db 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDYGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDYGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 2
US-10-303-161-10
; Sequence 10, Application US/10303161
; Publication No. US20030148459A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,161
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-161-10

Query Match 100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db 1 MSNNALVCGNPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIILKLDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIILKLDFFA 120
Qy 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Db 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDYGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDYGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 3
US-10-303-118-10
; Sequence 10, Application US/10303118
; Publication No. US20030157655A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,118
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-118-10

Query Match 100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db 1 MSNNALVCGNPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIILKLDFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIILKLDFFA 120
Qy 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Db 121 YIKYNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLKNCT 180
Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDYGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDYGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 4
US-10-303-128-10
; Sequence 10, Application US/10303128
; Publication No. US20030157656A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
US-10-303-128-10

```
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-128-10

Query Match      100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNINALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFFNPFVFFE 60
Db 1 MSNNINALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFFNPFVFFE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKKNLNKCT 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKKNLNKCT 120
Qy 121 YIKNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Qy 181 GFKNQKFKFINHSMACDQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 6
US-10-303-162-10
; Sequence 10, Application US/10303162
; Publication No. US20030157658A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-162-10

Query Match      100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNINALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFFNPFVFFE 60
Db 1 MSNNINALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFFNPFVFFE 60
Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKKNLNKCT 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSFGIDFYQDTNNLYAFDNNKKNLNKCT 120
Qy 121 YIKNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKNEIYNRQRTSGVTMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Qy 181 GFKNQKFKFINHSMACDQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Qy 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294
Db 241 INDILIPDKYAQERYYGKSRLENLHYKLIKDLIRLPSDIKHYLKEKYANKNR 294

RESULT 5
US-10-303-134-10
; Sequence 10, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-303-134-10

Query Match      100.0%; Score 1571; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNINALVCGNGPSLKNIDYKRLPKQFDVFRNCNQFYFEDRYFVGKDVKYVFFNPFVFFE 60
```


Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
DB 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDFFA 120

QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

QY 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
DB 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

QY 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294
DB 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294

RESULT 10
US-10-846-219-10
; Sequence 10, Application US/10846219
; Publication No. US20040219638A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/846,219
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-846-219-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
DB 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDFFA 120

QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

QY 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
DB 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Db 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
QY 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294
Db 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294

RESULT 11
US-10-821-604-10
; Sequence 10, Application US/10821604
; Publication No. US20040229263A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,604
; CURRENT FILING DATE: 2004-04-08
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-821-604-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60
DB 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQFYFEDRYFVGKDVYFNFNPFVFFE 60

QY 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDFFA 120
DB 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYSDFDAFLGHEIIKKLDFFA 120

QY 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
DB 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

QY 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
DB 181 GFKNQKFKFHNHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

QY 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294
DB 241 INDILIPDKYAQERYGKSRLENLHYKLIKDLIRLPSDIKHYLEKVKYANKNR 294

RESULT 12
US-10-847-983-10
; Sequence 10, Application US/10847983
; Publication No. US20040229272A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of

; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/847,983
; CURRENT FILING DATE: 2004-05-17
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-847-983-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

RESULT 13
US-10-821-573-10
; Sequence 10, Application US/10821573
; Publication No. US2004029313A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/821,573
; CURRENT FILING DATE: 2004-04-08
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: Gangliosides and Ganglioside Mimics
US-10-821-573-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

RESULT 14
US-10-850-807-10
; Sequence 10, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-850-807-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-821-573-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

RESULT 14
US-10-850-807-10
; Sequence 10, Application US/10850807
; Publication No. US20040259140A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,807
; CURRENT FILING DATE: 2004-05-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-850-807-10

Query Match 100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60
Db 1 MSNNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFEDRYFVGKDVYVFNPFVFFE 60

Qy 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120
Db 61 QYTSKKLIQNEEYNIENIVCSTINLEYIDGQFVDNFELYFSDAFLGHEIIKKLDKFFA 120

Qy 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db 121 YIKYNEIYNRQRTSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240
Db 181 GFKNQKFKFINHSMACDLQALDYLKRYDVNIYSLNSDEYFKLAPDIGSDFVLSKPKKY 240

Qy 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db 241 INDILIPDKYAORYYGGKSRLENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294

```
Db      121  YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Qy      181  GFKNQKFKFINHSMACDLQALDYLKRYDNYIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db      181  GFKNQKFKFINHSMACDLQALDYLKRYDNYIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Qy      241  INDILIPDKYAQERYGKSKRLKENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db      241  INDILIPDKYAQERYGKSKRLKENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
```

RESULT 15

```
US-10-850-125-10
; Sequence 10, Application US/10850125
; Publication No. US20040259203A1
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/850,125
; CURRENT FILING DATE: 2004-05-19
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (CstII) from C. jejuni strain NCTC 11168
US-10-850-125-10
```

```
Query Match      100.0%; Score 1571; DB 16; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.6e-125;
Matches 294; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60
Db      1  MSNNALVCGNGPSLKNIDYKRLPKQFDVFRNCQFYFEDRYFVGKDVKYVFNPFVFFE 60

Qy      61  QYVTSKKLIQNEEYNIENIVCSTINLEYIDGFQFVDNFELYFSDAPFLGHEIIKKLKDFFA 120
Db      61  QYVTSKKLIQNEEYNIENIVCSTINLEYIDGFQFVDNFELYFSDAPFLGHEIIKKLKDFFA 120

Qy      121  YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180
Db      121  YIKNEIYNQRITSGVYMCATAVALGYKSIYISGIDFYQDTNNLYAFDNNKKNLNKCT 180

Qy      181  GFKNQKFKFINHSMACDLQALDYLKRYDNYIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240
Db      181  GFKNQKFKFINHSMACDLQALDYLKRYDNYIYSLNSDEYFKLAPDIGSDFVLSKKPKKY 240

Qy      241  INDILIPDKYAQERYGKSKRLKENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
Db      241  INDILIPDKYAQERYGKSKRLKENLHYKLIKDLIRLPSDIKHYLEKRYANKNR 294
```

Search completed: June 13, 2005, 21:22:52
Job time : 158 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 16:47:14 ; Search time 4914 Seconds
(without alignments)
2899.030 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNIALVCGNGPSLKNIID.....IRLPDIKHYLKEKYANKNR 294

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
-Q=/cg2 1/USPTO.spool	-SUFFIX=rge -MINMATCH=0_1 -LOOPEXT=0
-DB=GenEmbl	-QFMT=fastap -TRANS=human40.cdi -LIST=45
UNITS=bits	-START=1 -END=1 -MATRIX=blosum62
-DOALIGN=200	-THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc	-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10735419	@CGN 1.1 4200 @runat 13062005 165724 12648 -NCPU=3
-NO MMAP	-LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120	-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7	-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```
-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cg2 1/USPTO.spool -SUFFIX=rge -MINMATCH=0_1 -LOOPEXT=0
-DB=GenEmbl -QFMT=fastap -TRANS=human40.cdi -LIST=45
UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10735419 @CGN 1.1 4200 @runat 13062005 165724 12648 -NCPU=3
-NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7
```

Database :

```
GenEmbl:*
1: gb.ba:*
2: gb.htg:*
3: gb.in:*
4: gb.em:*
5: gb.ov:*
6: gb.pat:*
7: gb.ph:*
8: gb.pl:*
9: gb.pr:*
10: gb.ro:*
11: gb.sts:*
12: gb.sy:*
13: gb.un:*
14: gb.vi:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1571	100.0	3071	1 AF195055	Campylobacter jejuni
2	1571	100.0	4668	1 AF257460	Campylobacter jejuni
3	1571	100.0	13484	1 AF400047	Campylobacter jejuni
4	1571	100.0	13484	1 AY044156	Campylobacter jejuni

5	1571	100.0	282183	1	CJ11168X4	Campylobacter jejuni
6	823	52.4	876	6	AF216647	Campylobacter jejuni
7	823	52.4	876	6	BD249791	Campylobacter jejuni
8	823	52.4	876	6	AR271700	Sequence
9	823	52.4	876	6	AR481782	Sequence
10	823	52.4	876	6	AR527381	Sequence
11	823	52.4	876	6	AX934425	Sequence
12	823	52.4	11474	1	AF130984	Campylobacter jejuni
13	823	52.4	11474	6	BD249790	Campylobacter jejuni
14	823	52.4	11474	6	AR271699	Sequence
15	823	52.4	11474	6	AR481781	Sequence
16	823	52.4	11474	6	AR527380	Sequence
17	823	52.4	11474	6	AX934424	Sequence
18	823	52.4	12576	1	AY297047	Campylobacter jejuni
19	817	52.0	876	6	BD249793	Campylobacter jejuni
20	817	52.0	876	6	AR271702	Sequence
21	817	52.0	876	6	AR481784	Sequence
22	817	52.0	876	6	AR527383	Sequence
23	817	52.0	876	6	AX934429	Sequence
24	816	51.9	11455	1	AY044868	Campylobacter jejuni
25	813	51.8	876	6	BD249792	Campylobacter jejuni
26	813	51.8	876	6	AR271701	Sequence
27	813	51.8	876	6	AR481783	Sequence
28	813	51.8	876	6	AR527382	Sequence
29	813	51.8	876	6	AX934427	Sequence
30	813	51.8	12388	1	AF401529	Campylobacter jejuni
31	812	51.7	11442	1	AF400048	Campylobacter jejuni
32	804	51.2	873	6	AR481786	Sequence
33	804	51.2	873	6	AR481787	Sequence
34	804	51.2	873	6	AR527385	Sequence
35	804	51.2	873	6	AR527386	Sequence
36	804	51.2	873	6	AX934434	Sequence
37	804	51.2	873	6	AX934436	Sequence
38	804	51.2	876	6	BD249794	Campylobacter jejuni
39	804	51.2	876	6	AR271703	Sequence
40	804	51.2	876	6	AR481785	Sequence
41	804	51.2	876	6	AR527384	Sequence
42	804	51.2	876	6	AX934431	Sequence
43	804	51.2	4749	1	AF305571	Campylobacter jejuni
44	804	51.2	11474	1	AF215659	Campylobacter jejuni
45	804	51.2	12390	1	AF401528	Campylobacter jejuni

ALIGNMENTS

RESULT 1	AF195055	Campylobacter jejuni strain MSC57360 hypothetical protein, SiatC (siatC), and SiatA (sitaA) genes, complete cds.	3071 bp	DNA	linear	BCT 06-NOV-2000
LOCUS	AF195055					
DEFINITION	AF195055					
ACCESSION	AF195055					
VERSION	AF195055.1	GI:11095586				
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						
source						
1..3071						
/organism="Campylobacter jejuni"						
/mol_type="genomic DNA"						
/strain="MSC57360"						
/db_xref="taxon:197"						

CDS	/note="serogroup: HS1"	61	GlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIleVal	80
	59. .943			
	/note="Ore1"	239	CATACTACTAGTAAAAAAGCTTATACAAAACGAAGAATATAAACATAGAAAATATAGTT	298
	/codon_start=1			
	/transl_table=11			
	/product="hypothetical protein"	81	CysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeu	100
	/protein_id="AAG29922.1"			
	/db_xref="GI:11095587"	299	TGTTCTACTATTAAATTTAGAAATACATAGATGGATTTCAATTTGTTGATATTTGAATTA	358
	/translations="MSMNALVCGNPSLKNIDYKRLPKQDFVFCNOFPYEDRYFV			
	GDKVYVFNPFPVFQYTSKGLIONEYNIENIVCSNTINLEYIDGFQFYPDPFELYP	101	TyrPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLeuLysAspPhePheAla	120
	SDAPLHEIILKLDPAFYIKYNEIINRQITSGVTCATAVAGLQYIISGIDFYQ			
	DTNLLAFDNKKNLNCKTGFKNQKPKFINHSMACDLQALDYLKRYDVIYSLNSD	359	TACTTTAGTCATGCTTTTTTAGTCATGAGATTATAAAAAGCTTAAAGATTTTTTGGCC	418
	EYFKAPDIDGSDVFLSKPKKYINDILIPDKYAQERYGKKSRLKENLHYLIKDLIR			
gene	LPSDIKHYLEKEYANKNR"	121	TyrIleLysTyrAsnGluIleTyrAsnArgGlnAtrGlnIleThrSerGlyValTyrMetCys	140
	924. .1955			
	/gene="siaC"	419	TATATTAAATATTAATGAATATACATAGCAGAAAGAAATTACTAGTGGAGTTTATATGTGT	478
CDS	924. .1955			
	/gene="siaC"	141	AlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGln	160
	/notes="static acid synthetase; required for sialylation of			
	LOS core"	479	GCTACAGCTGTGGCTCTGGGATATAAAGTATATATATAAGTGGTATTGATTTTATCAA	538
	/codon_start=1			
	/transl_table=11	161	AspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysLeuLeuAsnLysCysThr	180
	/product="SiaC"			
	/protein_id="AAG29923.1"	539	GATACAAATAATTTATATGCTTCGATAATAATAAAAAAATTTATTAATAATCAATGCAC	598
	/db_xref="GI:11095588"			
	/translations="MOIKDKLTISQKNPLIIPKIGINHGSLIAKLMVDAKRAGA	181	GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla	200
	KIHKOTHIVEDSMQEAKNVI PGNANISIVIEQCALNYKDELAKVEVKEQGLVY			
	LSTPFSRAANRLDMGVSAYKIGSGECNNYPLIKHIAQPKPKMILSTGMSIESIKP	599	GGTTTTAAAAATCAAAAATTTCAATTTCAATCAATGCGATGTCATTTTACAGCT	658
	TVKILRDYEPFVLLTNNLYPTPSHLVQLWLELYKEFNCLYGUSDHNTNNLCIG			
	AIAGASVIERHFTSMRKGPDIVCSMDSTLKDILINOTQEMVLLRGDNNKNPLKEE	201	LeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyr	220
	QVTDIFAFASVWSIKDKKEILSDMDNIWVRPSKGGISAKDFEALIGKRAKDKIKNN			
	IQTWDQDFE"	659	TTAGATTATCTAATGAAAAGGTATGATGTGAATATTTTATAGTTTAAATTCAGATGAGTAT	718
gene	1952. .3067			
	/gene="siaA"	221	PheLysLeuAlaPheProAspIleGlySerAspPheValLeuSerLysLysProLysTyr	240
CDS	1952. .3067			
	/gene="siaA"	719	TTTAAATTCGCTCCGGATATTGGAAGTGATTTGTATTGAGTAAAAAACCACAAAAATAT	778
	/notes="N-acetylmannosamine synthetase; required for			
	sialylation of LOS core"	241	IleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSer	260
	/codon_start=1			
	/transl_table=11	779	ATAAATGATATATTGATACCAATAGTATGACAGAAAGATATTTATGCAAAAAAAGT	838
	/product="SiaA"			
	/protein_id="AAG29924.1"	261	ArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIleArgLeuProSerAsp	280
	/db_xref="GI:11095589"			
	/translations="MKKIVFVSGTRADFSGIKSLMKMKNVSNFELFIFVTGMHMSKK	839	AGGTTAAAAAGAAAATCTCCATTATAAGCTAATAAAGATCTAATTAGACTTCCCAAGTGAC	898
	FSGTYMEIKGCFKNYKYNHDKYQMDKALSSITDGFSGFTHELEPDLIVVHGDRV			
	EPLAAIVGSLNNILVAHLEGCELSCTIDESLRHLSKLAHLHLVNDETAKRLIOMG			
	EDEKSIPIGSPDLELLNNTISLDEAKKYDIFKNYALAI FHPITTEINSIYKQSEE	281	IleLysHisTyrLeuLysGluLysTyrAlaAsnLysAsnArg	294
	FVNALIKSRNYIVIPNDLDGPELILQNYERLKNRERFKIPFSURFPFYFISLLKNAN			
	FIIGNSSCIIEALYLNINGILVGSQDQGTINDKTI RYNABEKOILEAILNTSKCTN	899	ATAAAACACTACTTAAAGGAAAAAATATGCAAAATAAAAAATAGA	940
	ITNKRLEILNSSEQFYRLKKNNLIFTINKQKIFMDKK"			
ORIGIN				
Alignment Scores:		4668 bp	DNA	linear
Pred. No.:	4.69e-116			BCT 03-DBS-2000
Score:	1571.00			Campylobacter jejuni 2,3-sialyl transferase (catIII), putative
Percent Similarity:	100.00%			sialic acid synthase (neuB1), N-acetyl mannosamine synthetis
Best Local Similarity:	100.00%			protein (neuCi), and beta 1,4-N-acetylglactosaminyltransferase/
Query Match:	100.00%			CMP-NeuNAC synthetase fusion protein (cgt/neuA1 fusion) genes,
DB:	1			complete cds.
		AF257460		
		AF257460.1	GI:11527837	
US-10-735-419-10 (1-294) x AF195055 (1-3071)				
QY	1	MetSerMetAsnIleAsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAsp	20	Campylobacter jejuni
DB	59	ATGAGTATGATATTAATGCTTTGGTATGCGGTAAATGGACCTAGTTTAAAAAATATAGAT	118	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
QY	21	TyrLysArgLeuProLysGlnPheAspValPheArgCysAsnGlnPheTyrPheGluAsp	40	Campylobacteraceae; Campylobacter.
DB	119	TATAAAGGTTACCAAAACAAATTTGATTTAGATGCAATCAGTTTATTTTGAAGAT	178	1 (bases 1 to 4668)
QY	41	ArgTyrPheValGlyLysAspValLysTyrValPhePheAsnProPheValPheGlu	60	Guerry,P., Ewing,C.P., Hickey,T.E., Prendergast,M.M. and Moran,A.P.
DB	179	AGATAATTTTGTGGGTAAAGATGAAATATGTTTTTTTTTAAATCCCTTTGTGTTTTTGA	238	Sialylation of lipooligosaccharide cores affects immunogenicity and
				serum resistance of Campylobacter jejuni
				Infect. Immun. 68 (12), 6656-6662 (2000)
				20536403
				11083778
				Guerry,P. and Ewing,C.P.

TITLE Direct Submission
JOURNAL Submitted (18-APR-2000) Enteric Diseases, Naval Medical Research Center, 503 Robert Grant Ave., Silver Spring, MD 20910, USA
FEATURES Location/Qualifiers
source 1. 4668
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="MSC 57360"
/serotype="HS01"
/db_xref="taxon:197"
53. .937
/notes="type strain of HS01 serotype"
53. .937
/genes="cstIII"
53. .937
/genes="cstIII"
/notes="cstIII"
/notes="Cst; involved in sialylation of LOS core"
/codon_start=1
/transl_table=11
/product="2,3-sialyl transferase"
/protein_id="AAG37020.1"
/db_xref="GI:11527838"
/translations="MSMNALVCGSPSLKNDYKRLPKQFDVFRCNQFYEDRYFV
GDKVYFNPFPVFFQYTSKLIQNEEYNIENIVCSITINLEYIDGFOFVDFNYLF
SDAPLGHETIKKLDPEAYIKWEIYNRQITSGVVMCAVALGKSVISIGIDYQ
DTNLYAFDNKNLKNLCKCTGFKNOKEKFNHSMACDLQALDYLKRYDNIYSLNSD
EYKLPADIGSDFFLSPKPKYINDILIPDKYAEYRYGKSKRLKENLHYKLKOLIR
LPSDIKHYLKEKTVANKR"
918. .1949
/genes="neuB1"
918. .1949
/genes="neuB1"
/notes="neuB1"
/codon_start=1
/evidence=not experimental
/transl_table=11
/product="putative sialic acid synthase"
/protein_id="AAG37021.1"
/db_xref="GI:11527839"
/translations="MQIKDKLTISQKNPLIIPFIEIGININGSLEIAKLMVDAAKAGA
KIIKHQTHIVEDMSQKRNVPGNANI SIYEIQCALNYKDEALKEVKEQGLVY
LSTPFGRAANRDMGVSAIKYIGECNNYPLIKHIAQFKPMIISTGMSIESIKP
TVKILRDYETPFVLLHTNLYPTPSHLVRLOAMLEYKBFNCLYGLSDHTNNLACIG
AIALGASVLERHETSMRKGPDIVCSMDSELTLDLINTQEMVLLRGDNKNPKLEE
QVTIDFAPASVSIKDIKGEILLMDNIWVRFSKGGISAKDFEALIGRAKDKIKNN
IQUTWDDFE"
1946. .3061
/genes="neuC1"
1946. .3061
/genes="neuC1"
/notes="neuC1"
/codon_start=1
/transl_table=11
/product="N-acetyl mannoseamine synthesis protein"
/protein_id="AAG37022.1"
/db_xref="GI:11527840"
/translations="MKKIVFVSQTRADFSKIKSLMKMKNSENEFELPIFVTGHRMSKK
FGSTYMEIKCGKPNKYKINHDKYYQMDKALSTTIDGFSKFTHIEPLDIVHGDV
EPLAAIVGSLNNILVAHIEGSGELSTIDESLHAIKSLAHILVNDEIAKRLIQMG
EDEKSIPIIGSPDLELNNITSLDEAKYVDIKFNKYATAIPIPIITTEINSYKQSEE
FVNALIKSEKNYIVIPNNDLGFELILQNVRLKNNRPFKIPSLRFEYFISLLKNN
FLIGNSSCIKEALYNINGILVGRQDGRDITKIRVNASEKDILEAILNTSKTN
ITNKRLEILNLSSEQFVRLKNNILFTINKQKIFMDKK"
3058. .4668
/genes="cgt/neuA1 fusion"
3058. .4668
/genes="cgt/neuA1 fusion"
/note="Cgt/NeuA1; apparent fusion of beta
1,4-N-acetylglactosaminyltransferase and CMP-NeuNAc
synthetase"
/codon_start=1
/transl_table=11
/product="beta 1,4-N-acetylglactosaminyltransferase/
CMP-NeuNAc synthetase fusion protein"

/protein_id="AAG37023.1"
/db_xref="GI:11527841"
/translations="MTLFYKIIAFURLKIDKKLPDNEYFLNLNKKIYNKHKHGFDP
FPNSKDTKSPINPWFPIRVKNEATLRVLSLEMLPAIQRGVIGYNDCTDGSSEIILE
FKQKPSFIPVKYPHEVQIENQSEENKLSHYNNYVASFIPQDEMLIKIDNVHYDAK
KLVKSPYMAKNTAVRPFRIFLIDKIVIONIGECGFIDGGDOLLIKQNSVPIERM
VSKSQMIDPEKTVKELYEQOILPKHILQAELOWHFPALKYHRNDYQKHLDALT
LEDFKIHYHRKIKKINTYMLDEKVIRESILDKFKLSKKMTLAIIPARAGSGIKNNK
NLALLHDPRLTYFINAANKSKYVDKIVLSSDDDDILEYGTQGVVDVLRKPKLEALDD
TTSDKVLVLTFFYKDYENIVLLQTSPLRTNVHIDEAFKFNKNSNALISVVECDN
KILKAPIDDNGNLKIGICDNKYPFMPROKLPKTYMSGAIYIVKSNLFLANNPTFLOEKT
SCVIMDEKASLDITDTEDLKRVNNISFL"
ORIGIN
Alignment Scores:
Pred. No.: 7.17e-116 Length: 4668
Score: 1571.00 Matches: 294
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-10-735-419-10 (1-294) x AF257460 (1-4668)
Qy 1 MetSerMetAsnIleAsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAsp 20
Db 53 ATGAGTATGAATATTAATGCTTTGGTATGCGGTAATGACCTAGTTTAAAAAATATAGAT 112
Qy 21 TyrLysArgLeuProLysGlnPheAspValPheArgCysAsnGlnPheTyrPheGluAsp 40
Db 113 TATAAAGGTACCAAAACAATTTGATGTATTAGATGCAATCAGTTTATTTTGAAGAT 172
Qy 41 ArgTyrPheValGlyLysAspValLysTyrValPhePheAsnProPheValPhePheGlu 60
Db 173 AGATATTTTGTGGTAAAGATGTAAATATGTTTTTTTAAATCCTTTTGTGTTTTTGAA 232
Qy 61 GlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIleVal 80
Db 233 CAATACCTACTAGTAGTAAAAAATTTATACAAACGAAAGATATATAACATAGAAAAATATAGT 292
Qy 81 CysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeu 100
Db 293 TGTTCCTACTATTAAATTTAGANTACATAGATGGATTTCAATTTTGTGTATTTTGAATTA 352
Qy 101 TyrPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLeuLysAspPhePheAla 120
Db 353 TACTTTAGTGTGCTTTTTTAGTGCATGAGATTATATAAAAGCTTAAAGATTTTTTTTGC 412
Qy 121 TyrIleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCys 140
Db 413 TATATTAAATATAATGAATAATATACATATGCGCAAGAAATTTACTAGTGGAGTTTATATGTGT 472
Qy 141 AlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGln 160
Db 473 GCTACAGCTGTGGCTCTGGGATATAAAGATATATATATAGTGGTATTGATTTTATCAA 532
Qy 161 AspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysLysLeuLeuAsnLysCysThr 180
Db 533 GATACAAATAATTTATATGCTTCGATAATAATAAAAAAAATTTATTAATAAATGCACCT 592
Qy 181 GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla 200
Db 593 GGTTTTAAAAAATCAAAAAATTTCAATTTTCAATTCATTCATTCATTCATTCATTCATTC 652
Qy 201 LeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyr 220
Db 653 TTAGATTATCTAATGAAAAGGTATGATGGAATATTATATTATTAATTTCAAGATGAGTAT 712
Qy 221 PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr 240
Db 713 TTTAAATTTGCTCCGGATATTGGAAGTATTTTGTATTGATTCAGTAAAAAATCAAAAAATAT 772
Qy 241 IleAsnAspIleLeuLeuProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSer 260

Db	773	ATAAATGATATATTGATACAGATAGTATGCACAGAAAGATATTATGAAAAAAGT	832
Qy	261	ArgLeuLysGluAsnLeuHisTyrLysLeuLeuLysAspLeuLeuArgLeuProSerAsp	280
Db	833	AGGTTAAAGAAATCTCCATTATAAGCTAATAAAGACTAATATTAGACTTCCAAGTGAC	892
Qy	281	IleLysHisTyrLeuLysGluLysTyrAlaAsnLysAsnArg	294
Db	893	ATAAACACTACTTAAGGAAATATGCAAATAAATAATAGA	934
RESULT 3			
AF400047			
LOCUS			
DEFINITION			
	AP400047	13484 bp	DNA linear BCT 03-JAN-2002
			Campylobacter jejuni heptosyltransferase I (waaC) gene, partial
			cds; lipid A biosynthesis acyltransferase (htrB), putative
			two-domain glycosyltransferase, putative glycosyltransferase,
			putative galactosyltransferase, beta-1,3-galactosyltransferase,
			alpha-2,3-sialyltransferase (cst-III), sialic acid synthase
			(neuB1), putative N-acetylglucosamine-6-phosphate 2-epimerase
			(neuC1), beta-1,4-N-acetylglucosaminyltransferase/CMP-Neu5Ac
			synthetase fusion protein (neuA), and putative glycosyltransferase
			genes, complete cds; heptosyltransferase II (waaF) gene, partial
			cds; and unknown genes.
ACCESSION	AF400047		
VERSION			
KEYWORDS			
SOURCE	AF400047.1	GI:15150361	
ORGANISM			
	Campylobacter jejuni		
	Campylobacter jejuni		
	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;		
	Campylobacteraceae; Campylobacter.		
REFERENCE	1	(bases 1 to 13484)	
AUTHORS	Gilbert M., Karwaski M.F., Bernatchez S., Young N.M., Taboada E.,		
	Michniewicz J., Cunningham A.M. and Wakarchuk W.W.		
TITLE	The Genetic Bases for the Variation in the Lipo-oligosaccharide of		
	the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF		
	STALYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE		
JOURNAL	J. Biol. Chem.	277 (1), 327-337	(2002)
PUBMED	11689567		
REFERENCE	2	(bases 1 to 13484)	
AUTHORS	Gilbert M., Michniewicz J., Karwaski M.-F., Cunningham A. and		
	Wakarchuk W.W.		
TITLE	Direct Submission		
	Submitted (14-JUL-2001) Institute for Biological Sciences, National		
JOURNAL	Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A		
	0R6, Canada		
FEATURES		Location/Qualifiers	
source	1.	13484	
	/organism="Campylobacter jejuni"		
	/mol_type="genomic DNA"		
	/strain="ATCC 43430"		
	/serotype="0:2"		
	/db_xref="ATCC:43430"		
	/db_xref="taxon:197"		
	/notes="Penner-type strain"		
	<1..348		
gene	/genes="waaC"		
CDS	<1..348		
	/gene="waaC"		
	/codon_start=1		
	/transl_table=11		
	/product="heptosyltransferase I"		
	/protein_id="AAK85412.1"		
	/db_xref="GI:15150362"		
	/translation="PAKEVNLNSGIDETHIELAPKNLSELMAFTOMNLLIINDSGP		
	THLAPALNKSITIFGAKTSYRNAFTHINKIIDAGKIQNAKHIDKDSDFCITRIEE		
	DIPKLAGLLNEK"		
	338..1225		
gene	/genes="htrB"		
CDS	338..1225		
	/gene="htrB"		
	/codon_start=1		
	/transl_table=11		

/product="lipid A biosynthesis acyltransferase"			
/protein_id="AAK85413.1"			
/db_xref="GI:15150363"			
/translation="WKNSDRIVLSLYILKFFVTWPECILHFLALIVARIAYFLNKK			
HRKLINTNQIPQPTQKERDKLSKIYENFAQGDICLQNTTKKELINKVNFIN			
ENFLDALANRPPIFTTAHYGNWEILSLAYAAKGAISIVGKGLKSNIMYELSSOR			
TOFDLELDKGGIRQMSALKRGRTGLITDQDCVENESVRLKPFNKVEVYQMGASL			
IAQRNALIIPVAVKBDGKFCIEFFKVKDSQNASLEULTLYAQSCSEMIKKRPWEY			
FFHRRFASYNKEIYKAK"			
1222..2769			
/note="orf3"			
/codon_start=1			
/transl_table=11			
/product="putative two-domain glycosyltransferase"			
/protein_id="AAK85414.1"			
/db_xref="GI:15150364"			
/translation="WNLKQISAIIVKNAKQITLLECLNSLKDQDEILLNSSDNTL			
NIANFRKDFANLHYHSAFIGALKNLALSAYAKNDWILSDADEVLNESCIKELKN			
LELOEDNIIALSRKNLYKGEMKACGMWPDYVVRIFNKNFTFNDNLVHESLVPNSA			
KKILKNGLRHYAFRDISHLIDKMOVYSLWAKONHKKSGLVKANLRAFMTFFRNYF			
LKNGFLYGYKGFIIISVCNALTGFYFKYKLYEQKOKPKTCALIITTYNKERKLVLVD			
SVKNLAFLEPNEVLIAADGSKEDTARLLEBYQKDFPCPLAGHMQDEGFPAAKRNVAL			
KASKSEYIILIDGMILEKDFVKNHLEFAKRVILQGSRVILNKSESEILKNNYSAL			
AFNKGFGNQRISFIKAKVYKFSKLTFRKPSQLVKGSCTNMSFYKSDFAEAGFN			
EKFIGWREDSEFVARFLFNDGVFKRLFNALAYHIYHEENSKNMLEINHQIYLETIK			
NEKTTWR"			
2773..3945			
/note="orf4"			
/codon_start=1			
/transl_table=11			
/product="putative glycosyltransferase"			
/protein_id="AAK85415.1"			
/db_xref="GI:15150365"			
/translation="MKYGVVPIYVNEKYLKECLDSVINQTYTNLEILLVNDGSTDE			
NSLNAKEYTLDKRILIFDKKNGGLSSARNVGIEYFSGEYTYIERKENELNVFKII			
GNPNQYIKIYKYNALASDKLEISKIDYIIFLDSNVMELNCIECVVRMQNDVLW			
FDHACTYDDGIEDKGCTRMNVNFTQECTITPRDYAKQAIKVGSRANSISNHDGKMTDF			
SFLQOLKLFINKINIEDIHFGWLPASADSIIYLPKRLYLCRLRANSISNHDKVKTK			
ANVSFYKDLVEFFGENAKENYLKAASRMITALELIEFFKDKQKNESQATKEAFLP			
FVVKALMIKKEKDPNLKKEKLPKIPQIQPIPYDLWKIWKQIKGILDKINPAK"			
complement (3929..4924)			
/note="orf5"			
/codon_start=1			
/transl_table=11			
/product="unknown"			
/protein_id="AAK85416.1"			
/db_xref="GI:15150366"			
/translation="MQIQNNSLIYNTLTKKLSSPIPKSTRRKLRNHHIOYKLEHPKV			
TNLYSNNYINPFEKGIPIHDFEKKHYFKNDKIIWQFYQGNQASPMIQCFNSVQS			
QMKDDYTIILDKDNIKDYLDPFPFVIEKLENNFFGEKTIITFFSLLRYCLLATYGGI			
WCDASIFLSSKIPSELCDKDPFTFERSKIKPSKEKPKFKISGYFSYGVFNWDDPKV			
KMLSFMIAKTNTNTFIOTLKDILMNYWKNQESSENNHYFILHIIPELLKEHPVNDIY			
KWMSDIECHLLQFSAKEPENSTLWEEIQKQSPFLHKLTHPKSIKKSMDIKIILQS"			
4982..6151			
/note="orf6"			
/codon_start=1			
/transl_table=11			
/product="putative galactosyltransferase"			
/protein_id="AAK85417.1"			
/db_xref="GI:15150367"			
/translation="MKTQGVPIYVNEKYLKECLDSVINQTYTNLEILLVNDGSTDE			
NSLNAKEYTLDKRILIFDKKNGGLSSARNVGIEYFSGEYKLNKTOITKENSILIEF			
NTEGNPYEILTYVYSYKAFNEODLTNTFTPIIDYIFLDSQDYWELNCIECVPRM			
DGVDVWDSIRYHDIKSYFKHSLKDINIKKSCRINPIEMLKLRONKIKDPFAFA			
KSGIDFDYIKKKMKFKAIPAEHDLFGILLFSQAKNIYVYVPFYVYTRIANSLTN			
QKKIKNLIPFKDIFIAFENATLAKYFKYVSWVETSELELVRVYFVYENHDKLSN			
LLKDTILFYIKNAFKIKKFDKPLCIKEKQLLKPYSIKIRYKVMKFLISLRN"			
complement (6711..7052)			
/note="premature stop codon due to deletion in			
homopolymeric tract"			
/codon_start=1			
/transl_table=11			
/product="beta-1,3-galactosyltransferase"			

/protein_id="AAK85418.1"
/db_xref="GI:15150368"
/translation="MSQISILPTYNVEKVIARALESCINQNTQKDIIEIIVVDGNDK
SIDIAKEYASKDRIKIHIHNEENKLLRAREYGAKVATSPYIMFLDSDDYELNACEE
CIRILDMGGGG"
7107. .7991
/gene="cst-III"
7107. .7991
/gene="cst-III"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAK85419.1"
/db_xref="GI:15150369"
/translation="MSMNALVCGNGPSLKNIDYKRLPKQDFVRCNQYFBRDVFV
GKDVYFFNPFFVEGYTSKLIQNEEYNIENIVCSTINLEYIDGFVDVDFELYF
SDAPLHEIIEIKLQDFPAYIKNEIYNRQRTISGVYMCATVALGYKQYIISGIDFYQ
DTNLYAFDNKKLNKCTGFKNQKFKFNHSMACDLOALDYLKRYDNIYSLNSD
EYFKLADIGSDVLSKKPKYINDILI PDKYAERYGKKSLKENLHVLIKDLIR
LPSDIKHLYLKEKANKR"
7972. .9003
/gene="neuB1"
7972. .9003
/gene="neuB1"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAK85420.1"
/db_xref="GI:15150370"
/translation="MQIKDKLITISOKNPLIIPKIGINHNHSLKIAKLMVDAARAGA
KIIKHOTHIVEDSMOAKNVIQGNANISIVEIMEOCALNYKDELALKEVEKQGLVY
LSTPFSAAANRLDMGVSAKYIGSGECNNYPLKIIHQAKPKMIISTGNNSTESIKP
TVKILRDYIEIPVLLHTNLYPTPSHLVRLOAMLELYKFNCLYSDHTTNLACIG
AIALGASVLRHRTDTMDRGPDIVCSMDSEITLKDLINQTOEMVLGRDNNKNPLKEE
QVTDIFAPASVSVSIKDIKKEGILLSMDNIWVKRPSKGISAKDPEALILGKRAKDKIKNN
IQLTWDDPE"
9000. .10115
/gene="neuC1"
9000. .10115
/gene="neuC1"
/codon_start=1
/transl_table=11
/product="putative N-acetylglucosamine-6-phosphate
2-epimerase"
/protein_id="AAK85421.1"
/db_xref="GI:15150371"
/translation="MKKIVFSGTRADFSKIKSLMMKVNSNPFELFVFTGMHMKK
FGTYMBIEKCGFNIVKYINHDKYQMDKALSTIDGFSKFIHEIEPDLIVVHGRV
EPLAAIVGSLNLVAHIEGSELGSDTIDSLRHAISKLAHILVNDKIAKKLIOMG
EDKSIPIGSDPLELNLNTISDEAKKYDIKFNVAIAIFHPITTEINSLYKQSEE
FVALIKSEKNYIVIPNNDLGFEILQLQNERKNNERFKIIPSLRPFYFISLLKNN
FIIGNSSCIIEALYLNINGILVSGRODGRDINKTIRVNAEBKDILEAILNTSKCTN
ITNKRLEILNSSSQFYRLKKNLIFTINKQKIFMDKK"
10112. .11722
/gene="neuA1"
10112. .11722
/gene="neuA1"
/codon_start=1
/transl_table=11
/product="beta-1,
4-N-acetylgalactosaminyltransferase/CMP-Neu5Ac synthetase
fusion protein"
/protein_id="AAK85422.1"
/db_xref="GI:15150372"
/translation="MTLFFKIIAFLRLKIDKKLPDNEVFLNLNKKIYDEKHGFPD
FDPSKDTKSPLPWAFIRVKNKATTLRVLSLMLPAIORGVIGYNDYTDGSEIILE
PCKQYSPFIPVKYPHEVQIENPQSEENKLSHYNNYASFPIDQEWLIKIDVDHYDAK

Alignment Scores:

Pred. No.: 2,09e-115 Length: 13484
Score: 1571.00 Matches: 294
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0
DB: 1 Gaps: 0
US-10-735-419-10 (1-294) x AF400047 (1-13484)
QY 1 MetSerMetAsnIleAsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAsp 20
DB 7107 ATGAGTATGATATTAATGCTTTGGTATGCGGTATGGACCTAGTTTAAAAATATATAGAT 7166
QY 21 TyrLysArgLeuProLysGlnPheAspValPheArgCysAsnGlnPheTyrPheGluAsp 40
DB 7167 TATAAAGGTTACCAAAACAATTTGATGATATTTAGATGCAATCAGTTTTTATTTTGAAGAT 7226
QY 41 ArgTyrPheValGlyLysAspValLysTyrValPhePheAsnProPheValPhePheGlu 60
DB 7227 AGATATTTTGTGGGTAAAGATGTAAATATGTTTTTTTAAATCCTTTTGTGTTTTTGAA 7286
QY 61 GlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIleVal 80
DB 7287 CATACTATACTAGTAAAAAACTTATACAAACGAGAGATATTAACATAGAAAATATAGTT 7346
QY 81 CysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeu 100
DB 7347 TGTCTACTATTAAATTTAGAAATACATAGATGGATTTTCAATTTGTTGATAATTTTGAATTA 7406
QY 101 TyrPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLysAspPhePheAla 120
DB 7407 TACTTTAGTGCATCTTTTGTAGTCAATGAGATTTATAAAAAGCTTTAAAGATTTTGTGCC 7466
QY 121 TyrIleLysTyrAsnGluIleTyrAsnAlaGlnArgIleThrSerGlyValTyrMetCys 140
DB 7467 TATATTAAATATTAATGAAATATACATAGGCAAGAAATTAATCTAGTGGAGTTTATATGTGT 7526
QY 141 AlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGln 160
DB 7527 GCTACAGCTGTGCTCTGGGATATAAAAGTATATATATAAGTGGTATTGATTTTATCAA 7586
QY 161 AspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThr 180
DB 7587 GATACAAATAAATTTATATGCTTTTCGATTAATAATAAAAAAATTTATTAATAATGCACT 7646
QY 181 GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla 200
DB 7647 GGTTTTAAAAATCAAAAATTTCAAAATTTATTCATCATTCAATGCGATGTGATTTACAAGCT 7706
QY 201 LeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyr 220
DB 7707 TTAGATTTATCTAATGAAAGGTATGATGTAATTTATATATTTTAAATTCAGATGAGTAT 7766
QY 221 PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr 240
DB 7767 TTTAAAATTGGCTCCGGATATTGGAAGTGATTTTGTATTGATGATAAAAAACCAAAAAATAT 7826
QY 241 IleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSer 260
DB 7827 ATAAATGATATATTGATACCAAGTAAGTATGCACAAGAAAGATATTATGGAAAAAAGT 7886
QY 261 ArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIleArgLeuProSerAsp 280
DB 7887 AGCTTAAAGAAAAATCTCCATTATAGCTAAATAAAGATCTAATTAGACTTCCCAAGTGAC 7946
QY 281 IleLysHisTyrLeuLysGluLysTyrAlaAsnLysAsnArg 294
DB 7947 ATAAACACTACTTAAAGGAAAAAATATGCAAAATAAAAAATAGA 7988
RESULT 4
LOCUS AY044156 13484 bp DNA linear BCT 03-JAN-2002
DEFINITION Campylobacter jejuni heptosyltransferase I (waaC) gene, partial cds; lipid A biosynthesis acyltransferase (htxB), putative two-domain glycosyltransferase, putative glycosyltransferase, putative galactosyltransferase, beta-1,3-galactosyltransferase, alpha-2,3-sialyltransferase (cst-III), sialic acid synthase

(neuB1), putative N-acetylglucosamine-6-phosphate 2-epimerase (neuC1), beta-1,4-N-acetylgalactosaminyltransferase/CMP-Neu5Ac synthetase fusion protein (neuA1), and putative glycosyltransferase (waaV) genes, complete cds; heptosyltransferase II (waaF) gene, partial cds; and unknown genes.

AY044156
AY044156.1 GI:14993916

Campylobacter jejuni
Campylobacter jejuni
Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales; Campylobacteraceae; Campylobacter.

1 (bases 1 to 13484)
Gilbert,M., Karwaski,M.F., Bernatchez,S., Young,N.M., Taboada,E., Michniewicz,J., Cunningham,A.M. and Wakarchuk,W.W.
The Genetic Bases for the Variation in the Lipo-oligosaccharide of the Mucosal Pathogen, Campylobacter jejuni. BIOSYNTHESIS OF STYLYLATED GANGLIOSIDE MIMICS IN THE CORE OLIGOSACCHARIDE

J. Biol. Chem. 277 (1), 327-337 (2002)
11689567

2 (bases 1 to 13484)
Gilbert,M., Michniewicz,J., Karwaski,M.-F., Cunningham,A. and Wakarchuk,W.W.
Direct Submission
Submitted (06-JUL-2001) Institute for Biological Sciences, National Research Council of Canada, 100 Sussex drive, Ottawa, Ontario K1A 0R6, Canada

Location/Qualifiers
1. .13484
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 43429"
/serotype="0:1"
/db_xref="ATCC:43429"
/db_xref="taxon:197"
<1. .348
/gene="waaC"
1. .348
/genes="waaC"
/codon_start=1
/transl_table=11
/product="heptosyltransferase I"
/protein_id="AAK73176.1"
/db_xref="GI:14993917"
/translations="FAKEVNLNSGIDETHIELAPKNLELMFAFTQMDDLIGNSGGP
THLAFALNKASITIFGATPSYRNAFQTHINKIIDTGKTIQNAKHIDKDFCITRIIEE
DIFLAKGLNER"
320. .1225
/genes="htrB"
320. .1225
/genes="htrB"
/codon_start=1
/transl_table=11
/product="Lipid A biosynthesis acyltransferase"
/protein_id="AAK73177.1"
/db_xref="GI:14993918"
/translations="MPKAYLMKNSDRIYLSLYILKFFVTFMPDCILHLFALIVARIA
FHLNKKRIINTNLQICPPQYTKERDKLSKIYENAFQFGIDCLQONTMTKEILN
KYNFINELIDALNRPILITTAHYGNWEILSLAYAAKYGAISIVGKLSEINYE
ILSQSTQDIELDKGGIRQMLSKKRGTRILGILTDQCVENESVRLKPFNKEVNY
QMGASIIAQRSNLIIPVYAYKEDGKFCIEFFKVKDSQNASUEELTYQAQSCHEMIK
KRPWEYFFHRRFASYNKEIYKDAK"
1222. .2769
/codon_start=1
/transl_table=11
/product="putative two-domain glycosyltransferase"
/protein_id="AAK73178.1"
/db_xref="GI:14993919"
/translations="MNLKQISAIIVKNAQTILCECLNSIKDQDEIILLNBSNDNL
NANEPKQFANLHIHVSAPFIGGALKNLALSVAKNDDWILSIDADEVLENECIKELKN
LEIQEDNIIALSLENLKGEWIKACCGWPDVYVRIENKNFTFNDNLVHESVLPSNA
KKLYLKNGLRHAFRILSHLIDKQYSSIMAKQNIHKSGVLKANLRAFWTFRRNYF
LKNGLFYGYKFIISVCNALGTGPFKWKUYELQKQPKPTCALIITITYNOKERLKLVLID

SVKNLAFPLNEVLIAADSGKEDTARLIEEYQKDFPCPLKHIWQEDEGFAAKSRNNAI
KASKSYIILIDDKMILKDFVKNHLEPAKRVILQSGVILNKMSSEETLKNNTSL
ANKKGKRNQISIFIAKCVYKSKLTKEFKKSQLVKGSKTCNMSFYKSDFAIEGPN
EXFICWGRDSESVARPLFNDGVFKLKNALAYHIYHEENSKMLEINHQIYLETIK
NKTITWR"

2773. .3945
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAK73179.1"
/db_xref="GI:14993920"
/translation="MKKGVGVIPYVNEKYLKECLDSVINQTYTNLEILVNDGSTDE
NSLNTAKETLKDRIILFDKNGGLSSARNVGIYFSGEYETQYTERENELNVFKII
GNPNQIYKIYKQNALSDKLEISKIDYIIFLSDSNYWELNCIECVVMQNDVLM
FPHACIYDGDIGKQKTMNVNFTECTITPRDYAKQAIKVGSRDIFSWMGMDIF
SELKOLKAFINKINEDIHFGWVLPASADSIYLPKRLYLCELTRANSISNHDKKVTK
ANVSFYKDIYEFFGGENAKQNYLKAASRMITALELIBEFFQDKQKNENSQAIEKALP
FYVKALMIKYPFKOPLNIKEKLPKPIQTKIPYDLWKIWKIQIKGILDKINFAK"
complement (3929. .4924)
/notes="orf5"
/codon_start=1
/transl_table=11
/product="unknown"
/protein_id="AAK73180.1"
/db_xref="GI:14993921"
/translation="MQIOQNNSLIYNTLTKLSSPIKSTRRKLNHHIOYKLEHPKV
TYLNNYINPFLGKIHFDKFKHYFKNDKIIMQFVQGNQKQSPMIQQCFNSVQS
QMKDYYTIIILDKDNIKYLDPPFPVIEKLENNPFGEKITTFESDLRLVCLLATYGGI
WDSATIFLSSKIPSELCKDQDFTFERSKPKSEKELKFKIGSYGIFGVNDDDFKV
KMSDFMAKTNTTFTQTLKDILMNYWNEQSENHYFILHIFELKEHGFVNDIY
KNMSDIECHLQFSAREKFTNWBEEIQSQSLHLKTHPKSIKQSDMIDKIIIQS"
4982. .6151
/notes="orf6"
/codon_start=1
/transl_table=11
/product="putative galactosyltransferase"
/protein_id="AAK73181.1"
/db_xref="GI:14993922"
/translation="MKTGVGVIPYVNEKYLKECLDSVINQTYTNLEILVNDGSTDE
NSLNTAKETLKDRIILFDKNGGLSSARNVGIYFSGEYKLNKQTOIKNESLIEF
NIEGNPYEITYYKSYKAPNEQDLTFTYFIIDYIIFLSDSDWELNCIECVPRM
DGVVVFSDSIBYDIEKSYFHSRLKDNINKKCRINPILBWLKLRQNKIKDPAFA
WSGIIDFYIDKDKMKFDKDAIFAEHDLFGILLFSQAKNIYVYKPYRIRANSLTN
QDKITKONILYFKDIFIAFENATLAKYEFKYVSWVETSLELVRFYENYHDKKISS
LKUDTILFYIKNAFKIKKFDKPLCIKEKLQLLKPYKISKIRYKVMKFLISLN"
complement (6711. .7052)
/notes="premature translation stop due to deletion in
homopolymeric tract on the complementary strand"
/codon_start=1
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAK73182.1"
/db_xref="GI:14993923"
/translation="MSQSIILPTVYVNEKYIARALESCINQTFKQIDBIILVDDCGNDK
SIDIAKYASKDDRIKIIHNEENKLLRAREYGAKVATSPYIMFLSDSDYLELNACEE
CIKILDMGGG"
7107. .7991
/genes="cst-III"
7107. .7991
/genes="cst-III"
/codon_start=1
/transl_table=11
/product="alpha-2,3-sialyltransferase"
/protein_id="AAK73183.1"
/db_xref="GI:14993924"
/translation="MSMNIHALVCGNPSLKNIDYKRLPKQFDVFRCNQFYEDRYFV
GDKVYFNFNPFVPEQYTSKKLIINEEYNIENIVCSTINLEYIDGFOFVDFNFBLYF
SDAFIGHIIEKLKDFYAYIKYNEIYRQIRISGVYMCATVALGYSYISGIDIFYQ
FTNLNLYAFDNKKNLNLKCTGFKQKFPFINHSMACDLQALDLMKRYDNVYISLNSD
EYFKLAPDIGSDVLSKKPKPKYINDILIPDKYAQERYYGGKSRSLKNLHYKLIKDLIR
I3SDIKHYLKEKYANKR"
7972. .9003
/gene="neuB1"

CDS	7972..9003	/gene="neuB1"	7287	CAATACTACTAGTAAAAAACTTATACAAACGAGAAATATAACATAGAAAATATAGTT	7346	
	/codon_start=1		QY	81	CysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeu	100
	/transl_table=11		Db	7347	TGTTCTACTATTAAATTTAGAAATACATAGATGGATTTCAATTTTGTGATAAATTTTGAATTA	7406
	/product="sialic acid synthase"		QY	101	TyrPheSerAspAlaPheLeuGlyHisGluIleIleLeuLe	

sequencing team, Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA E-mail: parkhill@sanger.ac.uk
Notes:
Details of C. jejuni sequencing at the Sanger Centre are available on the World Wide Web.
(URL, http://www.sanger.ac.uk/Projects/C_jejuni/).

COMMENT

FEATURES

source

1. .282183
Location/Qualifiers
/organism="Campylobacter jejuni subsp. jejuni NCTC 11168"
/mol_type="genomic DNA"
/strain="NCTC 11168"
/db_xref="taxon:192222"
complement (43. .1443)
/gene="Cj1009c"
complement (43. .1443)
/gene="Cj1009c"
/notes="Cj1009c, unknown, len: 466 aa; 29.1% identity to HP0282"
/codon_start=1
/transl_table=11
/product="hypothetical protein Cj1009c"
/protein_id="CAB73265.1"
/db_xref="GI:6968445"
/db_xref="GOA:Q9PNT1"
/translation="MNILIIIDGILAKHFLERLCPEKGLGYFFTVVCSKKNLNI
SSEYIDLHYDPDTSSTARLEINMSDKQAFIVMQDEFETKKSVEALRSQNPNEIIM
DFWGLSVNDTHANLADARWTLSPRREMDLPDIALTAQYICLGVEIWEVKIPAGSIFA
YRHISIQKQWRVILLYRNSKIYFVKPSFVLEPNDISILIVGDPVVLQIFHNIRGKA
GQPFMPFGSNFALIDMKNMNQMRVLDITLUKTOKNNAKRFPIHVINPKLGVMYE
KLKJSDKEGVEFFDFYDFDKQISMLQNDIGLVVDIKNPEKEKQAFDDIKPIIM
KVGASDFDELKEAIIISADESELENNANVITDLSKQDFGVLIIYYNPNNSQNTDME
YFRLSKLYDKNIQIKNKNDENPLNLQYREDLLQFVSQKELLNRDFAFNLSLNLNR
HYFKMQNTQLFIPVE"
1473. .2594
/genes="tgt"
1473. .2594
/genes="tgt"
/EC_number="2.4.2.29"
/notes="Cj1010, tgt, probable queueine
tRNA-ribosyltransferase, len: 373 aa, similar to many e.g.
TGT_ECOLI queueine tRNA-ribosyltransferase (EC 2.4.2.29)
(375 aa), fasta scores; opt: 1016 z-score: 1190.9 E(): 0,
41.3% identity in 373 aa overlap. 56.7% identity to
HP0281"
/codon_start=1
/transl_table=11
/product="queueine tRNA-ribosyltransferase"
/protein_id="CAB73266.1"
/db_xref="GI:6968446"
/db_xref="GOA:Q9PNT0"
/db_xref="UniProt/Swiss-Prot:Q9PNT0"
/translation="MBFKLKHKGDMARVCEITTAHSTFLTPVPMVGTGVAKVSLDAN
DMKNEJDAKILLANTYHMLRPTSKVVKQFGGLHGTGKDFSLTDSGGQFAPLSUKN
SGHFNIGFKSHIDSRHLFTPQSLVDYQDFNSDIMLDDLVALPATKGRKVISV
DRTILWAKAEITHHKMNQKIGIQNIIFGIIQGGDYBERKRCALSLNMFDPGLAI
GGLSGEENALYPTQNLNPLNDYRPLMGVGTPEDLVENVRGVDMDFCMPTFR
NARNGTFFTSFGKFNFKAEFINDEHVIDSTCSYTCRNFSGRYLNLHFKAKELTFR
LASLNLHYLELARKREALNLNSFTQKRNPHYLRGK"
2596. .3363
/genes="Cj1011"
2596. .3363
/genes="Cj1011"
/notes="Cj1011, probable membrane protein, len: 255 aa; no
Hp match"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAB73267.1"
/db_xref="GI:6968447"
/db_xref="GOA:Q9PNS9"
/translation="MLELHENLKKILOAKNLETFYSEIYGQKIFVYVGLNLETLWLPND

gene

CDS

1473. .2594
/genes="tgt"
1473. .2594
/genes="tgt"
/EC_number="2.4.2.29"
/notes="Cj1010, tgt, probable queueine
tRNA-ribosyltransferase, len: 373 aa, similar to many e.g.
TGT_ECOLI queueine tRNA-ribosyltransferase (EC 2.4.2.29)
(375 aa), fasta scores; opt: 1016 z-score: 1190.9 E(): 0,
41.3% identity in 373 aa overlap. 56.7% identity to
HP0281"
/codon_start=1
/transl_table=11
/product="queueine tRNA-ribosyltransferase"
/protein_id="CAB73266.1"
/db_xref="GI:6968446"
/db_xref="GOA:Q9PNT0"
/db_xref="UniProt/Swiss-Prot:Q9PNT0"
/translation="MBFKLKHKGDMARVCEITTAHSTFLTPVPMVGTGVAKVSLDAN
DMKNEJDAKILLANTYHMLRPTSKVVKQFGGLHGTGKDFSLTDSGGQFAPLSUKN
SGHFNIGFKSHIDSRHLFTPQSLVDYQDFNSDIMLDDLVALPATKGRKVISV
DRTILWAKAEITHHKMNQKIGIQNIIFGIIQGGDYBERKRCALSLNMFDPGLAI
GGLSGEENALYPTQNLNPLNDYRPLMGVGTPEDLVENVRGVDMDFCMPTFR
NARNGTFFTSFGKFNFKAEFINDEHVIDSTCSYTCRNFSGRYLNLHFKAKELTFR
LASLNLHYLELARKREALNLNSFTQKRNPHYLRGK"
2596. .3363
/genes="Cj1011"
2596. .3363
/genes="Cj1011"
/notes="Cj1011, probable membrane protein, len: 255 aa; no
Hp match"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAB73267.1"
/db_xref="GI:6968447"
/db_xref="GOA:Q9PNS9"
/translation="MLELHENLKKILOAKNLETFYSEIYGQKIFVYVGLNLETLWLPND

gene

CDS

2596. .3363
/genes="Cj1011"
2596. .3363
/genes="Cj1011"
/notes="Cj1011, probable membrane protein, len: 255 aa; no
Hp match"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAB73267.1"
/db_xref="GI:6968447"
/db_xref="GOA:Q9PNS9"
/translation="MLELHENLKKILOAKNLETFYSEIYGQKIFVYVGLNLETLWLPND

gene

tRNA

3416. .3492
/gene="tRNA-Arg"
3416. .3492
/gene="tRNA-Arg"
/product="tRNA-Arg"
/note="tRNA Arg anticodon GCG, Cove score 83.29"
complement (3495. .3944)
/gene="Cj1012c"
complement (3495. .3944)
/gene="Cj1012c"
/note="Cj1012c, possible membrane protein, len: 149 aa; no
Hp match"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAB73268.1"
/db_xref="GI:6968448"
/db_xref="UniProt/TREMBL:Q9PNS9"
/translation="WVIFCAALALLFLGVIAYLITSDGKTIKKQTSQKQHVAKET
KKFDTDDRMIIAASDVKLTDIELKELAKLYVQTHKLGSKTSKLEDAAKKKLEFVSA
LAANTNASQTYSLNKKELKKIYSGKEIDAYEHMGLAKRKIKEDK"
complement (3961. .7206)
/gene="Cj1013c"
complement (3961. .7206)
/gene="Cj1013c"
/note="Cj1013c, probable membrane protein, len: 1081 aa;
contains three domains; aa 1-90 contains three membrane
spanning domains; aa 90-780 is non-membrane, and aa
780-1081 contains ten possible membrane spanning domains.
Some similarity in C-terminal membrane domain to eukaryotic
proteins e.g. CCSA_CHURE cytochrome C biogenesis protein
CCSA (353 aa), fasta scores; opt: 459 z-score: 515.4 E():
2.1e-21, 36.0% identity in 267 aa overlap, and to
prokaryotic equivalents e.g. CCMF_ECOLI cytochrome C-type
biogenesis protein CCMF (647 aa); BLASTP scores; E = 0.20,
24% identity in 293 aa overlap. 39.0% identity to HP0378
(domains 2+3 only)"
/codon_start=1
/transl_table=11
/product="putative membrane protein"
/protein_id="CAB73269.1"
/db_xref="GI:6968449"
/db_xref="GOA:Q9PNS7"
/db_xref="UniProt/TREMBL:Q9PNS7"
/translation="MKNIIKSGDLRVSVFLFLFALFALCATTFIESAYGTPTAMAMV
YTFWFVEYIQLLGINLLCGMPRYMFGKLPMLFIHISFLFILVGSAMTRYAGFEG
ILPIREHTONSLIESKTSLRISAIKDGERSYAVNDRYIGNLPFANSFKLNLGDDQ
AVLYKDLILNAHYTYKENNSDPLLVLMQSGQGVVDVKEGVEKNIGVNPAPFM
NDNVKAPFKVIDENITLSSSENHLFLMLDQCQLDKIGEKANERRLYEINDISFV
VKAASQAALLEGNSRPDESFLWFKSAWLEGRKTMISTFGFPQMNKNSLLHLFK
DFALNENKLETSYKNAKLSYKNESEFYIPEYKNPIMIELADQKFFISWALS
EOLPDIYLRDVLDRPGMSGPASYASEITVNNNNENFDYRI FNNVLDYDGYRYQ
SSYDDEKGTLSVNKDPKIPYIYIGFLLCGMFNLPHSRFTLRNLNKOTLTK
HNSVITIFLLSFGSEKTPAODINSTLPVWNTNHAALATLIVKQSGADGRMVPDTLS
REILEKHQSDSYKQNSNAVMSMLVDVQWQLEPFIILMPQONQAVRDALNLEIPS
AKYIAKOPFDENNRIRKQYVENANRNKRNAPGVDFKEIKIDENANVNLVFSSEL
FKFIPVQNNPNVNLAPPSAVTTLKGDEGHIVLALIQNYFSAVENAKDGNWIRADEG
LKFPIKEYEKIGYKMPSTKVEIMEIFSNKABIFVKLAPVLIAGLLILVFSKRVV
PNLKSFTFKVYVNLVLAFTHTVGLIRAYLSHAPMSGNGYESMVIYAWALSIGI
FYSRKSPTALSITSILSGVLVMVAHLSENNPOITNLVPLVNSYVSIHVSVITASYGF
LGLCALLGIFTFLMCFLLKDGKQKYNLNLNTEATINEMAMIFGLCIVTQNFILGA
IWANESWGRYSWDSKETWALVSILVYAAIILHRLMIPKYNQCFVPAWNSMPAYWVILM
TYFGVNYFLTGLHSAAGAAQIPNYVYVYMGFALMVVLAFFARRKKNFVGLK."
complement (7224. .7919)
/gene="livP"
complement (7224. .7919)
/gene="livP"
/note="Cj1014c, livP, probable branched-chain amino-acid
ABC transport system ATP-binding protein, len: 231 aa;

gene

CDS

7224. .7919
/gene="livP"
complement (7224. .7919)
/gene="livP"
/note="Cj1014c, livP, probable branched-chain amino-acid
ABC transport system ATP-binding protein, len: 231 aa;

similar to e.g. LIVE_ECOLI high-affinity branched-chain amino acid transport protein (237 aa), fasta scores; opt: 686 z-score: 797.6 E(: 0, 46.4% identity in 233 aa overlap. No Hp ortholog. Contains PS00017 ATP/GTP-binding site motif A (p-loop), PS00211 ABC transporters family signature, and Pfam match to entry PF00005 ABC tran"

/codon_start=1
/transl_table=11
/product="branched-chain amino-acid ABC transport system ATP-binding protein"
/protein_id="CAB73270.1"
/db_xref="GI:6968450"
/db_xref="COA:Q9PNS6"
/db_xref="UniProt/TrEMBL:Q9PNS6"
/translation="MLIVDLHVVYGLIEAVKIDPKFVETGHIVSLIGSNGACKTSTL NALLNSVKTGEYNYFLGYDTKRLHTLTVQKGLIALVPEGRVRFINLSVBNLKGAFN NAANYHLRBYMKLPRLASKKHALAGTSLSGEAOMLAISRALMSEPKLLMDPSL GLAPKIVGEVFDIIIVRLKEGIIILLVEQNAYSALKISDVAVYLENGRIVMENEAKNL IGDEIRKKYLGL"
complement(7299..7841)
/gene="livp"
/notes="Pfam match to entry PF00005 ABC tran, ABC transporters, score 147.40, E-value 2.5e-40"
complement(7482..7526)
/gene="livp"
/notes="PS00211 ABC transporters family signature"
complement(7797..7820)
/gene="livp"

misc_feature
4.51e-114 Length: 282183
1571.00 Matches: 294
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 1

misc_feature
US-10-735-419-10 (1-294) x CJ11168X4 (1-282183)

QY 1 MetSerMetAsnIleAenAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAsp 20
Db 134466 ATGAGTATGAATTAATGCTTTGGTATGCGGTAAATGGACTAGTTAAAAAATATAGAT 134525
QY 21 TyLysArgLeuProLysGlnPheAspValPheArgCysAsnGlnPheTyPheGluAsp 40
Db 134526 TATAAAGGTTACCAACAATTTGATGTTATAGTGAATCAGTTTATTTTGAAGAT 134585
QY 41 ArgTyPheValGlyLysAspValLysTyPheValPhePheAsnProPheValPheGlu 60
Db 134586 AGATATTTTGTGGGTAAAGATGTAAATATATGTTTTTTTAAATCCTTTTGTGTTTGTAA 134645
QY 61 GlnTyTyThrSerLysLysLeuIleGlnAsnGluGluTyPheAsnIleGluAsnIleVal 80
Db 134646 CATACTATCTAGTAAAAAATCTTATACAAACGAGNATATACATAGAAAATATAGTT 134705
QY 81 CysSerThrIleAsnLeuGluTyPheIleAspGlyPheGlnPheValAspAsnPheGluLeu 100
Db 134706 TGTTCTACTATTAATTTAGAAATACATAGATGGATTTCAATTTGTTGATAATTTTGAATTA 134765
QY 101 TyPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLysLysAspPhePheAla 120
Db 134766 TACTTTAGTGTGCTTTTTTAGGTGATGAGATTATAAAAAGCTTAAAGATTTTTTTTGGC 134825
QY 121 TyPheLysTyPheAsnGluIleTyPheAsnArgGlnArgIleThrSerGlyValTyMetCys 140
Db 134826 TATATTAATATATATGAAATATACATAGGCAAGAATTTACTAGTGAGTTTATATGTGT 134885
QY 141 AlaThrAlaValAlaLeuGlyTyPheLysSerIleTyPheSerGlyIleAspPheTyGln 160
Db 134886 GCTACAGCTGTGGCTCGGCATATAAAAGTATATATAAGTGTGATTGATTTTATCAA 134945
QY 161 AspThrAsnAsnLeuTyPheAlaPheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThr 180

Db 134946 GATACAAATAATTTATATGCTTCGATAATAATAATAAAAAAATTTATTAATAATGCAC 135005
QY 181 GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla 200
Db 135006 GGTTTTAAAAATCAAAAAATTCAAATTTATCAATCATTTCAATGGCATGTGATTTACAAGCT 135065
QY 201 LeuAspTyPheLeuMetLysArgTyPheValAsnIleTyPheSerLeuAsnSerAspGluTyP 220
Db 135066 TTAGATTATCTAATGAAAAAGGTATGATGCAATATTTATATAGTTTAAATTCAGATGAGTAT 135125
QY 221 PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysLys 240
Db 135126 TTTAAATTTGGTCCGATATTGGAAGTGATTTTGATTTGATGATTAATAAACCACAAAAAATAT 135185
QY 241 IleAsnAspIleLeuIleProAspLysTyPheAlaGlnGluArgTyPheTyPheLysLys 260
Db 135186 ATAAATGATATATTGATACAGATAAGTATGCACAAGAAAGATATTTATGCAAAAAAAGT 135245
QY 261 ArgLeuLysGlnAsnLeuHisTyPheLysLeuIleLysAspLeuIleArgLeuProSerAsp 280
Db 135246 AGGTTAAAAAGAAATCTCCATTATAAGCTAATAAAAAAGTCTAATTAGACTTCCAAGTGAC 135305
QY 281 IleLysHisTyPheLysGlnLysTyPheAlaAsnLysAsnArg 294
Db 135306 ATAAAAACACTACTTTAAGGAAAAAATATGCAATTAATAAATAGA 135347

RESULT 6
AF216647
LOCUS
DEFINITION
AF216647
VERSION
AF216647.1 GI:17221830
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1..876
/organism="Campylobacter jejuni"
/mol_type="genomic DNA"
/strain="ATCC 700297"
/serotype="ATCC 700297"
/db_xref="Penneer 0:19"
/db_xref="ATCC:700297"
/db_xref="taxon:197"
1..876
/gene="cst-II"
1..876
/gene="cst-II"
/note="bifunctional sialyltransferase"
/codon_start=1
/transl_table=11
/product="alpha-2,3/8-sialyltransferase"
/protein_id="AAL36462.1"
/db_xref="GI:17221831"
/translation="MKKVIAGNGPSLKEIDYSLRPNDFVFCNDFYEDKYYLGKK CKAFVFNPLFFFEQYVTLKHLIQNQEYTELIMCSNQAHLENENFVTFYDFPDA HLYGDFPKOLKPNAYFKPHEIYFNORITSGVVMCAVALGKYEIYLSGIDFYQGS SVAFDTKOKNLKLPNFKNDSHVTHGSKNTDVKALEFLEKTYKIKLYCLCPNSLLA NFIELAPNLNSFIQEKNNYTKDILIPSEAYGKFSKNFKIKENIYYKLIKD

ORIGIN		LLRLPSDIKHFKGK"	
Alignment Scores:		6.22e-57	Length: 876
Pred. No.:	823.00	Matches:	161
Score:	823.00	Conservative:	43
Percent Similarity:	69.62%	Mismatches:	71
Best Local Similarity:	54.95%	Indels:	18
Query Match:	52.39%	Gaps:	6
DB:	1		
US-10-735-419-10 (1-294) x AF216647 (1-876)			
QY	8	LeuValCysGlyAsnGlyProSerLeuLeuAsnIleAspTyrLysArgLeuProLysGln	27
Db	13	ATTATTGCTGCAATGACCAAGTTTAAAGAAATGGATTATTCAAGACTACCAATGAT	72
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
Db	73	TTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA	132
QY	48	ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys	67
Db	133	TGCAAGGCAGTATTTTACAACTCTTCTTTTGAACATATCTACTTAAACAT	192
QY	68	LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
Db	193	TTAAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTCTTAATTACACCAAGCT	252
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107
Db	253	CATCTAGAAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCTGATGCTCATTTG	312
QY	108	GlyHisGluIleIleLysLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle	127
Db	313	GGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAAT	372
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147
FEATURES			
source			
1..876			
/organism='Campylobacter jejuni'			
/mol_type='genomic DNA'			
/db_xref='taxon:197'			
ORIGIN			
Alignment Scores:			
Pred. No.:	6.22e-57	Length:	876
Score:	823.00	Matches:	161
Percent Similarity:	69.62%	Conservative:	43
Best Local Similarity:	54.95%	Mismatches:	71
Query Match:	52.39%	Indels:	18
DB:	6	Gaps:	6
US-10-735-419-10 (1-294) x BD249791 (1-876)			
QY	8	LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27
Db	13	ATTATTGCTGGAATGGACCAAGTTTAAAGAAATGGATTATTCAAGACTACCAATGAT	72
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
Db	73	TTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA	132
QY	48	ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys	67
Db	133	TGCAAGGCAGTATTTTACAACTCTTCTTTTGAACATATCTACTTAAACAT	192
QY	68	LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
Db	193	TTAAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTCTTAATTACACCAAGCT	252
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107
Db	253	CATCTAGAAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCTGATGCTCATTTG	312
QY	108	GlyHisGluIleIleLysLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle	127
Db	313	GGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAAT	372
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147

RESULT 7	BD249791	876 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Campylobacter glycolyltransferases for biosynthesis of				
DEFINITION	gangliosides and ganglioside mimics.				
ACCESSION	BD249791	1	GI:33059561		
VERSION	JP 2002535992-A/2.				
KEYWORDS	Campylobacter jejuni				
SOURCE	Campylobacter jejuni				
ORGANISM	Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;				
REFERENCE	Campylobacteraceae; Campylobacter.				
AUTHORS	1 (bases 1 to 876)				
TITLE	Gilbert, M. and Wakarchuk, W.W.				
JOURNAL	Campylobacter glycolyltransferases for biosynthesis of				
COMMENT	gangliosides and ganglioside mimics				
	Patent: JP 2002535992-A 2 29-OCT-2002;				
	NATIONAL RESEARCH COUNCIL OF CANADA				
	OS	Campylobacter	jejunii		
	PN	JP 2002535992-A/2			
	PD	29-OCT-2002			
	PP	01-FEB-2000	JP 2000597438		
	PR	01-FEB-1999	US 60/118213, 31-JAN-2000	US	06/495406 PI
	MICHEL GILBERT, WARREN W WAKARCHUK				
	PC	C12N15/09, C12N1/21, C12N9/10, C12N9/88, C12N15/00	CC		
	Campylobacter glycolyltransferases for biosynthesis of				
	gangliosides and				
	CC	ganglioside mimics			
	FH	Key	Location/Qualifiers		
	FT	source	1..876		
	FT	/organism='Campylobacter jejuni'.			
FEATURES					
source	1..876				
	/organism='Campylobacter jejuni'				
	/mol_type='genomic DNA'				
	/db_xref='taxon:197'				
ORIGIN					
Alignment Scores:					
Pred. No.:	6.22e-57	Length:	876		
Score:	823.00	Matches:	161		
Percent Similarity:	69.62%	Conservative:	43		
Best Local Similarity:	54.95%	Mismatches:	71		
Query Match:	52.39%	Indels:	18		
DB:	6	Gaps:	6		
US-10-735-419-10 (1-294) x BD249791 (1-876)					
QY	8	LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27		
Db	13	ATTATTGCTGGAATGGACCAAGTTTAAAGAAATGGATTATTCAAGACTACCAATGAT	72		
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47		
Db	73	TTTGATGTATTAGATGTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA	132		
QY	48	ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys	67		
Db	133	TGCAAGGCAGTATTTTACAACTCTTCTTTTGAACATATCTACTTAAACAT	192		
QY	68	LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87		
Db	193	TTAAATCCAAAATCAAGAATATGAGACCGAACTAAATATGTGTCTTAATTACACCAAGCT	252		
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107		
Db	253	CATCTAGAAAATGAAAATTTTGTAAAACTTTTACGATTTATTTCTGATGCTCATTTG	312		
QY	108	GlyHisGluIleIleLysLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle	127		
Db	313	GGATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTTACGAAAT	372		
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147		


```

Db      373  TATTTCAATCAAGAAATTAACCTCAGGGGTTTATATGTCGAGTAGCCATAGCCCTAGGA 432
Qy      148  TylysseriletyrileSerGlyleAspPheTyrglnAspThrAsnAsnLeuTyAla 167
Db      433  TACAAGAAATTTATCTTCGGAATTTGATTTTATCAAAATGGGTCTATCT---TATGCT 489
Qy      168  PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490  TTTGATACTAAACAAAAATCTTTTAAATTTGGCTCCTTAATTTTAAATAATGATAAATCA 549
Qy      188  LysPheileAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550  CACTATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAAACT 609
Qy      208  TyrAspValAsniletyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610  TACAAAAATAAACTATATTTGCTTATGCTTAACAGCTCTTTTAGCAAAATTTTCAAAATACTA 669
Qy      224  AlaProaspiledGlySerAspPheValLeuSerLysLysProLysLysTyAlaAsnAsp 243
Db      670  GCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
Qy      244  IleLeuileProaspLysTyAlaGlnGluArgTyrTyGly----- 257
Db      727  ATACTCATACCTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAATAATTT 774
Qy      258  -----LysLysSerArgLeuLysGluAsnLeuHisTyLysLeuLysAspLeuile 275
Db      775  AATTTTAAAAATAAAATAAAGAAATATTTATTACAAGTTGATAAAGATCTATTA 834
Qy      276  ArgLeuProSerAspLysLysHisTyLysLeuLysGluLys 288
Db      835  AGATTACCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 8
AR271700
LOCUS      AR271700
DEFINITION Sequence 2 from patent US 6503744.
ACCESSION AR271700
VERSION    AR271700.1 GI:29703245
KEYWORDS   'Unknown.'
ORGANISM   'Unknown.'
REFERENCE  1 (bases 1 to 876)
AUTHORS    Gilbert,M. and Wakarchuk,W.W.
TITLE      Campylobacter glycosyltransferases for biosynthesis of gangliosides
           and ganglioside mimics
JOURNAL    Patent: US 6503744-A 2 07-JAN-2003;
FEATURES   Location/Qualifiers
           source
             1..876
             /organism="unknown"
             /mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.:      6.22e-57      Length:      876
Score:          823.00      Matches:    161
Percent Similarity: 69.62%      Conservative: 43
Best Local Similarity: 54.95%      Mismatches: 71
Query Match:    52.39%      Indels:    18
DB:              6      Gaps:      6

US-10-735-419-10 (1-294) x AR271700 (1-876)

Qy      8  LeuValCysGlyAsnGlyProSerLeuLysAsnLysAspTyrLysArgLeuProLysGln 27
Db      13  ATTATTGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTCAAGACTACCAAAATGAT 72
Qy      28  PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73  TTTGATGATTATAGATGATATCAATTTTATTTTGAAGATTAATCTATCTTGGTAAAAA 132

```

```

Qy      48  ValLysTyValPhePheAsnProPheValPhePheGluGlnTyrTyThrSerLysLys 67
Db      133  TGCAGGCGAGTATTTTACAAATCTCTATTTCTTTTGAACAATACTACACTTTTAAAAACAT 192
Qy      68  LeuileGlnAsnGluGluTyrAsnLysGluAsnLysValCysSerThrLysAsnLeuGlu 87
Db      193  TTAATCCAAATCAAGATATAGACGGAACATAATATATGTTCTTAATATTACACCAAGCT 252
Qy      88  TyrileAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253  CATCTAGAAATCAAAATTTTGTAAACCTTTTACGATTATTTTCTCGATGCTCATTTG 312
Qy      108  GlyHisGluileleLysLysLysAspPhePheAlaTyrLysLysTyAsnGluile 127
Db      313  GGAATATGATTTTTCACAACTTAAAGATTTTAAATGCTTATTTTAAATTTTACGAAAT 372
Qy      128  TyrAsnArgGlnArgileThrSerGlyValTyMetCysAlaThrAlaValAlaLeuGly 147
Db      373  TATTTCAATCAAGAAATTAACCTCAGGGGTTTATATGTCGAGTAGCCATAGCCCTAGGA 432
Qy      148  TyLysSeriletyrileSerGlyleAspPheTyrglnAspThrAsnAsnLeuTyAla 167
Db      433  TACAAAGAAATTTATCTTCGGAATTTGATTTTATCAAAATGGGTCTATCT---TATGCT 489
Qy      168  PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490  TTTGATACTAAACAAAAATCTTTTAAATTTGGCTCCTTAATTTTAAATAATGATAAATCA 549
Qy      188  LysPheileAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550  CACTATATCGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAAACT 609
Qy      208  TyrAspValAsniletyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610  TACAAAAATAAACTATATTTGCTTATGCTTAACAGCTCTTTTAGCAAAATTTTATAGAACTA 669
Qy      224  AlaProaspiledGlySerAspPheValLeuSerLysLysProLysLysTyAlaAsnAsp 243
Db      670  GCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
Qy      244  IleLeuileProaspLysTyAlaGlnGluArgTyrTyGly----- 257
Db      727  ATACTCATACCTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAATAATTT 774
Qy      258  -----LysLysSerArgLeuLysGluAsnLeuHisTyLysLeuLysAspLeuile 275
Db      775  AATTTTAAAAATAAAATAAAGAAATATTTATTACAAGTTGATAAAGATCTATTA 834
Qy      276  ArgLeuProSerAspLysLysHisTyLysLeuLysGluLys 288
Db      835  AGATTACCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 9
AR481782
LOCUS      AR481782
DEFINITION Sequence 2 from patent US 6699705.
ACCESSION AR481782
VERSION    AR481782.1 GI:47243417
KEYWORDS   'Unknown.'
ORGANISM   'Unknown.'
REFERENCE  1 (bases 1 to 876)
AUTHORS    Gilbert,M. and Wakarchuk,W.W.
TITLE      Campylobacter glycosyltransferases for biosynthesis of gangliosides
           and ganglioside mimics
JOURNAL    Patent: US 6699705-A 2 02-MAR-2004;
FEATURES   Location/Qualifiers
           source
             1..876
             /organism="unknown"
             /mol_type="genomic DNA"

ORIGIN

```

Alignment Scores:		6.22e-57	Length:	876
Pred. No.:	823.00	Matches:	161	
Score:	69.62%	Conservative:	43	
Best Local Similarity:	54.95%	Mismatches:	71	
Query Match:	52.39%	Indels:	18	
DB:	6	Gaps:	6	
US-10-735-419-10 (1-294) x AR481782 (1-876)				
QY	8	LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27	
Db	13	ATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCACAGACTACCAAAATGAT	72	
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47	
Db	73	TTTGATGATTTAGATGTAATCAATTTTGAAGATAAATCTATCTGTGTAATAAAA	132	
QY	48	ValLysTyrValPhePheAsnPropheValPheGluGlnTyrTyrThrSerLysLys	67	
Db	133	TGCAAGGCAGTATTTTACAATCTTCTTTTGAACAATCTACACTTTTAAACAT	192	
QY	68	LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87	
Db	193	TTAATCCAAATCAAGAATATGAGCCGAACCTAAATATGTTCTTAATACACCAAGCT	252	
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107	
Db	253	CATCTAGAAAATGAAAATTTGTAAACCTTTTACGATATTTCTGATGCTCATTTG	312	
QY	108	GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle	127	
Db	313	GGATATGATTTTTCACCAACCTTAAAGATTTTAAATGCTTATTTTAAATTCACGAAAT	372	
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147	
Db	373	TATTTCAATCAAGAATATACCTCAGGGTTTATATGTGTGCAGTAGCCATAGCCCTAGGA	432	
QY	148	TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyrAla	167	
Db	433	TACAAAGAATTTATCTTCGGAAATGATTTTATCAAAATGGGTCATCT---TATGCT	489	
QY	168	PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe	187	
Db	490	TTTGATACTAAACAAAATCTTTTAAATTTGGCTCCTTAATTTTAAATATGATAATTCA	549	
QY	188	LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg	207	
Db	550	CACATATCGGCACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTGAAAAAAT	609	
QY	208	TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu	223	
Db	610	TACAAATAAATACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT	669	
QY	224	AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp	243	

LOCUS AR527381
DEFINITION Sequence 2 from patent US 6723545.
ACCESSION AR527381
VERSION AR527381.1 GI:53914402
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 876)
AUTHORS Gilbert,M. and Wakarchuk,W.W.
TITLE Polypeptides having .beta.-1,4-GalNAC transferase activity
JOURNAL Patent: US 6723545-A 2 20-APR-2004;
FEATURES
Location/Qualifiers
1..876
source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 6.22e-57 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 6
US-10-735-419-10 (1-294) x AR527381 (1-876)
QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCACAGACTACCAAAATGAT 72
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGATTTAGATGTAATCAATTTTGAAGATAAATCTATCTGTGTAATAAAA 132
QY 48 ValLysTyrValPhePheAsnPropheValPheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCAAGGCAGTATTTTACAATCTTCTTTTGAACAATCTACACTTTTAAACAT 192
QY 68 LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAATCCAAATCAAGAATATGAGCCGAACCTAAATATGTTCTTAATACACCAAGCT 252
QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAAATGAAAATTTGTAAACCTTTTACGATATTTCTGATGCTCATTTG 312
QY 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGATATGATTTTTCACCAACCTTAAAGATTTTAAATGCTTATTTTAAATTCACGAAAT 372
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAATATACCTCAGGGTTTATATGTGTGCAGTAGCCATAGCCCTAGGA 432
QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyrAla 167
Db 433 TACAAAGAATTTATCTTCGGAAATGATTTTATCAAAATGGGTCATCT---TATGCT 489
QY 168 PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTGATACTAAACAAAATCTTTTAAATTTGGCTCCTTAATTTTAAATATGATAATTCA 549
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGGCACATAGTAAATAACAGATATAAAGCTTTTGAATTTCTGAAAAAAT 609
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAATACTATATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTATGCTTAT 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 GCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAAAAATTTTCAAAAAATATTT 774
QY 258 -----LysLysSerArgLysGlnAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db 775 AATTTTAAAAATAAATAAATAAAGAAATATATTTTATACAGTTGATAAAGATCTATTTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873
RESULT 10
AR527381

```

Db      670 GCGCCAAATTAATTCATATTTATCATACAGAAAAA---ATAACTACACTAAGAT 726
Qy      244 ILLeuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTAGGCT-----TATGCAAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLeuLysAspLeuLe 275
Db      775 AATTTTAAAAAATAAATAAAGAAATATTATTATCAAGTTGATATAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspLeuLysHisTyrLeuLysGluLys 288
Db      835 AGATTACCTAGTATATAAGCAATTATTCAAGGAAAA 873

RESULT 11
AX934425
LOCUS      Sequence 2 from Patent WO02074942.
DEFINITION
ACCESSION AX934425
VERSION    AX934425.1 GI:40641672
KEYWORDS
SOURCE
ORGANISM   Campylobacter jejuni
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
            Campylobacteraceae; Campylobacter.

REFERENCE
1 Gilbert,M. and Wakarchuk,W.W.
  Campylobacter glycosyltransferases for biosynthesis of gangliosides
  and ganglioside mimics
  Patent: WO 02074942-A.2 26-SEP-2002;
  National Research Council of Canada (CA)
FEATURES
source     1..876
            /organism="Campylobacter jejuni"
            /mol_type="unassigned DNA"
            /db_xref="taxon:197"
            1..876
            /note="unnamed protein product; bifunctional
            alpha-2,3/alpha 2,8-sialyltransferase Campylobacter
            sialyltransferase II (cstII) from C. jejuni strain OH4384
            (Osp 7a of lipooligosaccharide (LOS) biosynthesis locus)"
            /codon_start=1
            /transl_table=11
            /protein_id="CAF04166.1"
            /db_xref="GI:40641673"

CDS
            1..876
            /note="translation="MKVVIAGNGPSLKEIDYSLRPNDFVFRCNQFYFEDKYLGGK
            CKAVFNPILFFQYVTLKHLIONQVETELIMCSNVNOAHLENENFVXTFYDYPDA
            HLGYDFKQLKFNAPKPEIYFNQRIISGVYMCVAIALGYKEIYLSGIDPYQNGS
            SYAFDTKQMLLKLPNFRNDNSHIIGHSKNTDIKALEFLEKTYIKLKYCLCPNSLLA
            NFTELAPNLSNFIIQEKNNYTKDILIPSEAYGKPSKNIKFKIKENIYYKLIKD
            LLRLPSDIKHYPKKG"

ORIGIN
Alignment Scores:
Pred. No.:      6.22e-57      Length:      876
Score:          823.00      Matches:    161
Percent Similarity: 69.62%      Conservative: 43
Best Local Similarity: 54.95%      Mismatches: 71
Query Match:    52.39%      Indels:    18
DB:              6          Gaps:      6

US-10-735-419-10 (1-294) x AX934425 (1-876)

Qy      8 LeuValCysGlyAenGlyProSerLeuLysAsnLeuAspTyrLysArgLeuProLysGln 27
Db      13 ATATTGCTGGAATGACCAAGTTTAAAGAAATTTGATTATTCAAGACTACCAATGAT 72
Qy      28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTTGATGATTAGATGAATCAATTTTATTTGAAGATAAATATCTATCTTGTGTAAGAAA 132
Qy      48 ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrSerLysLys 67

```

```

Db      133 TGCAGGCGAGTATTTTACAACTCTATCTCTTTTGTGACAACTACACTTAAAACAT 192
Qy      68 LeuileGlnAenGluGlyTyrAenileGluAenileValCysSerThrileAenLeuGlu 87
Db      193 TTAATCCAAATCAAGATATGAGACCGAACTAATATTATGTGTCTTAATTACAACCAAGCT 252
Qy      88 TyrileAspGlyPheGlnPheValAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAATGAAATTTTGTAAACACTTTTACGATTTATTTCTGTGTCATTTG 312
Qy      108 GlyHisGluileileLysLysLeuLysAspPheAlaTyrileLysTyrAenGluile 127
Db      313 GGATATGATTTTTCACAACTTAAAGATTTTAATGCTTATTTTAAATTTTACGAAT 372
Qy      128 TyrAsnArgGlnArgileThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAAATTAACCTCAGGGTTTATATGTGTGACGAGCCATAGCCTAGGA 432
Qy      148 TyrLysSerileTyrileSerGlyileAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCCGGGAATTTGATTTTATCAAAATGGGTCTCTCT--TATGCT 489
Qy      168 PheAspAsnAsnLysLysAenLeuLysCysThrGlyPheLysAenGlnLysPhe 187
Db      490 TTTGATACTAAACAAAAAATCTTTTAAATTTGGCTCTTAATTTTAAATATGATATTCA 549
Qy      188 LysPheileAenHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGGCATAGTAAATAACAGATATATAAAGCTTTTACGAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAATATAAACTATATTGCTTATGTCCTCAACAGCTTTTACCAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrileAenAsp 243
Db      670 GCGCCAAATTTAAATTCAAATTTTATCATACAGAAAAA---ATAACTACACTAAGAT 726
Qy      244 IleLeuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTAGGCT-----TATGCAAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAenLeuHisTyrLysLeuLeuLysAspLeuile 275
Db      775 AATTTTAAAAAATAAATAAAGAAATATTATTATCAAGTTGATATAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspLeuLysHisTyrLeuLysGluLys 288
Db      835 AGATTACCTAGTATATAAGCAATTATTTCAAAGGAAAA 873

RESULT 12
AF130984
LOCUS      11474 bp      DNA      linear      BCT 09-FEB-2000
DEFINITION Campylobacter jejuni lipooligosaccharide biosynthesis locus,
            partial sequence.
ACCESSION AF130984
VERSION    AF130984.1 GI:6940827
KEYWORDS
SOURCE     Campylobacter jejuni
            Campylobacter jejuni
            Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
            Campylobacteraceae; Campylobacter.
            1 (bases 1 to 11474)
            Gilbert,M., Brisson,J.R., Karwaski,M.F., Michniewicz,J.,
            Cunningham,A.M., Wu,Y., Young,N.M. and Wakarchuk,W.W.
            Biosynthesis of ganglioside mimics in Campylobacter jejuni OH4384.
            Identification of the glycosyltransferase genes, enzymatic
            synthesis of model compounds, and characterization of nanomole
            amounts by 600-mhz (1h) and (13)c NMR analysis
            J. Biol. Chem. 275 (6), 3896-3906 (2000)
            20127862
            PUBMED 10660542

```

```

REFERENCE 2 (bases 1 to 11474)
AUTHORS Gilbert M., Michniewicz, J. and Wakarchuk, W. W.
TITLE Direct Submission
JOURNAL Submitted (24-FEB-1999) Institute for Biological Sciences, National
Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario K1A
0R6, Canada

FEATURES             Location/Qualifiers
     source            1..11474
                        /organism="Campylobacter jejuni"
                        /mol_type="genomic DNA"
                        /strain="OH4384"
                        /db_xref="taxon:197"
                        <1..360
     CDS               /note="Helicobacter pylori RfaC homolog"
                        /codon_start=1
                        /transl_table=11
                        /product="heptosyltransferase I"
                        /protein_id="AAF31765.1"
                        /db_xref="GI:6940828"
                        /translations="KEYEFAKEVNLNSGIDETHIELAPKFNLEELMAFTKMDLIIGN
                        DSGPHTLAPALNKASITIFGATPSYRNAFQTHINKIIDTGKKIQNAKHIDKSDFCITR
                        IREEDIFKLAKGLNEK"
     CDS               350..1237
                        /note="Helicobacter pylori WaaM homolog"
                        /codon_start=1
                        /transl_table=11
                        /product="lipid A biosynthesis acyltransferase"
                        /protein_id="AAF31766.1"
                        /db_xref="GI:6940829"
                        /translations="MKNSDRIYLSLYILKFFVTFMPDCILHLALIVARIAFLNKK
                        HRKINTALQKPIPTQYKDERKSLKIYENFAQGIDCLQNTTKERILNKNVFN
                        ENFLIDALAKRIPIFTTAHYGNWEILSLAYAGAIATVGGKLSVEMVYILQSQR
                        TQFDILIDKQKGIQMLSALKKERALGILTDQCVENESVRLKFNKEVNYQMGASL
                        IAQRNALIPVAYEGGKFCIEFFKAKDSQNASLEELTLVQAQSCSEMIKKRPWEY
                        PPHRRPASYNEBIYKAGK"
     CDS               1234..2490
                        /note="Neisseria meningitidis LgtF homolog"
                        /codon_start=1
                        /transl_table=11
                        /product="putative glycosyltransferase"
                        /protein_id="AAF31767.1"
                        /db_xref="GI:6940830"
                        /translations="MNLKQISVIIIVKNAEQTLLCCLNSLKDFDEIILLNNESSDNTL
                        KIANEFKDFANLIYIYHNAFIGFGALKNLNSYAKNDWILSIDADEVLENECIEKLKN
                        LKQEDNIITALKSRNKIDYKAWICGWPYVVLRI FNKNFTFNDNLVHESLVLPNSA
                        KTYLKNGLKHSYKDI SHLIDKMOVYSSIMAKQNLHKKSGVLKANLRAFTWFFRNYF
                        LKNGFLYGGKPTIISVCSALGTFPKYMKLYELQRKPTCALIITYNQKRLKLVLD
                        VKNLAFLEVLIAADGSKEDTARLIEYQKDFPCLPKHIWQEDGFKLSKSRNKTII
                        KNADSYIIVIDGDMILEKDFITKEHLFEAQRKLFQGSRVILNKKESEBELNKDDYRI
                        IFNKKDFKSSKNSFLAKIFYSLSKKR"
     CDS               2786..3955
                        /note="similar to Streptococcus pneumoniae Cps14J"
                        /codon_start=1
                        /transl_table=11
                        /product="putative glycosyltransferase"
                        /protein_id="AAF31768.1"
                        /db_xref="GI:6940831"
                        /translations="MKKIGWPIYVNVKYLRECLDSVINQTYTNLEIILVNDGSTDE
                        HSLNTAKETLAKDKRTLFDKNGGLSSARNIGIYFSGEYKLNKTQHKIENSLIEP
                        QLDGNPNYNIYKAYSSQAFNNEKDLTNFTYSDIIFLDSNTYKMLNCIEECVIRM
                        KQVDVLWFDHCDTIEDNIKNKKHKTMEIFDPFKECIITPKPYANRALSVGRSDISFG
                        WNGMTIDFNLFQIKLQIFINFI NEDIHFGIILFASANKIYVLSQKLYLCRLRANSISN
                        HDKMTIKNVSVYFKDIYETFGENAKEAKNYLKAASRVITALKLIEFFFKQKNENALA
                        IKETFLPCYAKKALMIKKPKDPLNKEQLVLVLIKPIQTKLPYDIWKFWQIKNI"
     gene              4025..5068
                        /gene="cgtA"
     CDS               4025..5068
                        /gene="cgtA"
                        /note="glycosyltransferase"
                        /codon_start=1
                        /evidence=experimental
                        /transl_table=11
                        /product="beta-1,4-N-acetylgalactosaminyltransferase"

/protein_id="AAF31769.1"
/db_xref="GI:6940832"
/translations="MLFQSYFYKICLPIFPKRIRHKIKKTFLLKNIQRDKIDSYLPK
KTLVKNYNNEDLIKNAKAIIGEGHGYFNDEKSKPKSPLNPAFIRVKNEAITL
KASLSLPAIQRGVIYNDCTDSEETILEFCQKYPSPFIPIKYPIEIQNPKSEEN
KLYSYNVASPIPKDWEMLIKIDVDHIIYDAKELKYSFYIPKQKYDVVSYSRVDIHYFN
DNFFLCKNNNGILKEPGDCLINNYLNAKWEVLIDRIINNWKATKQSFSSNIHSLE
QLYKXRIILFHTLANNYHFFLKGRHQADIIKYNNWISIEFKKFFYLQNIHNHKIEPSMI
SKETLKKIFLTLP"
complement (5054..5959)
/gene="cgtB"
complement (5054..5959)
/gene="cgtB"
/note="glycosyltransferase"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="beta-1,3-galactosyltransferase"
/protein_id="AAF31770.1"
/db_xref="GI:6940833"
/translations="MEKISILILPTYNVEQYIARAJESCINOTFKDIEIIVDDCGNDN
SINAKESKDKRIKIHNKNGLLRAREYGVKVPSPYIMFLDPDPILELNACEE
CIKILDEQDEVDLVFNAINESNVSISYKFPDNGSFYSKKEFKVKIIAKKNLYTWMMG
KLIRKLYLEAFASLEKEDVKINMAEDVLLYPLMSQAQKIAYMNCNLYHYVPNNNS
ICNTKNEVLVKNNOIQLQLVNLNRYLQNYILNKYCSVLYLKYLYIYIYKIKRTKLM
VTLAKINILTLKILFKYKFKLQK"
6048..6923
/gene="cst-II"
6048..6923
/gene="cst-II"
/note="bifunctional sialyltransferase"
/codon_start=1
/evidence=experimental
/transl_table=11
/product="alpha-2,3/8-sialyltransferase"
/protein_id="AAF31771.1"
/db_xref="GI:6940834"
/translations="MKKVIIAGNGPSLKEIDYSRLPNDPVPFCNQFYPEDKYVLGKK
CKAFVNPILPPEQYVTLKHLIQOEYETELJMCNVAQAHLENFVKTFFVDPDA
HLGYDFQLQKDFNAYFPHEIYFNQRTISGVYMCALVATGKYIYLSGIDFYQNGS
HFIADTKQNLKULAPNFKNDNSHYIHSKNTDIKALEFLEKTYIKLYKLCPLCSLLA
NFIELAPNLNSNFIIOEKNNYTKDILIPSEAYGKPSKNINPKKI KENIYIYKLID
LRLPESDIKHYFKGK"
6924..7964
/note="Neisseria meningitidis SiaC homolog"
/codon_start=1
/transl_table=11
/product="sialic acid synthase"
/protein_id="AAF31772.1"
/db_xref="GI:6940835"
/translations="MKEIKIQNIITSEKAPLVVPEIGINHNHNGSLAKIMVDAAPST
GAKLIHQTHIVEDEMSKAAKVIFGNAKISIEIEMQKCALDYKDELAKYETKLG
VLTSPFRAGANRLDEMGVAFKFGSGECNNYPLIKHIAFKPKPMI VSTGNWSIESI
KPTVKILLDNEIPIFVMTHTNLPTPHNLVRLNMLLELKBEFCVWGLSDHTDNLAC
LGAVALGACVLERHFTDSMRSGPDIVCSMDTQALKELIIOEQMAIMRGNNSKAA
KOEQVITDFAFASVVSIIKDIKKEVLSDMDNIWVKRPLGGISAAEFENILGKALRDI
ENDTQLSYEDFA"
8021..9079
/note="involved in sialic acid biosynthesis; Neisseria
meningitidis SiaA homolog"
/codon_start=1
/transl_table=11
/product="SiaA-like protein"
/protein_id="AAF31773.1"
/db_xref="GI:6940836"
/translations="MYRVONSSEFELYIFATGMHLSKNFGYTVTKLYKNGKFNIEYFI
NYDKYFTDKALATTIDGFSRYVNELKPDLI VWHGDRIEPLAAAVIGALNLTVAHIE
GGEISGTTDDSLRAHAIKSLAHLVNDFAKRLMQLGEDEKSI FIIIGSPDLELDN
KISLNEAKKYVDINVENYALAMFHPVTTEITSIKNOADNLVKALI QSNKNYIYIPNN
DLGPELLILOSVEELKNRPFKLPFSLRFEYFITLLKNADFIIGNSSCILKEALYKTA
GILVSGRONGRIGNENTLUKRVANSDEILKAINTIHKKQDLFSKLEILDOSSKLFPEYL
QSGEFKLTQKVPKDIK"
9076..9741

```

/note="Haemophilus ducreyi Neua homolog"
/codon_start=1
/transl_table=11
/product="CMP-sialic acid synthetase"
/protein_id="AAF31774.1"
/db_xref="GI:6940837"
/translation="MSAIIIPARGSGKIKNNKLVLLNNKPLIYYTIIKAALNTKTSISK
VYVSSDEFLNTAKSNQNDILKRPISLAQDNTSDKVLHAKFYKDYEDVVFLQPT
SPKRNTHIDEAFNLKYNANALISVSECNDILKAFVCNEYGDLAGICNDSPYMP
ROKLPTKYMNGAIYILKIEFLNPNPSFLQSKTKHFLMDESSSLDIDCLEDLKKAQOI
WKK"
CDS
7929..10562
/note="similar to Rhodobacter capsulatus hypothetical
protein encoded by GenBank Accession Number AF010496"
/codon_start=1
/transl_table=11
/product="putative acetyltransferase"
/protein_id="AAF31775.1"
/db_xref="GI:6940838"
/translation="MEKITIKCNKNILNLKQNYIYTKTYENPRFSLRKTQDFTTF
FNFQHQSVAGLIGIEYCAKFSNIIHEDMSFSGSFLPHYTKVGRYSIGDSVSM
PLENQPMDRISTASTFTYVNHSDINACQHNKTPFIVNHNPSSITLHIIQDDVMI
GKDVLLKQGITLGTGCVIGORAVVTKDVPYALVAGIPAKIIKYRDEKTIERLLKIQ
WKKHFAFDYDIDLNLKINQYLDLLEKIIKKSISYYNPKUYFRDILELKSXKIFN
F"
CDS
complement(10554..11366)
/note="similar to Haemophilus influenzae hypothetical
protein H10868 encoded by GenBank Accession Number U32768"
/codon_start=1
/transl_table=11
/product="putative glycosyltransferase"
/protein_id="AAF31776.1"
/db_xref="GI:6940839"
/translation="MPQLSIIIPLPSNCSDFISRALQSCINQLKDIILLIIDDKSKDN
F"
Alignment Scores:
Pred. No.: 8,95e-56 Length: 11474
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 1 Gaps: 6

US-10-735-419-10 (1-294) x AF130984 (1-11474)

Qy	8	LeuValCysGlyAenGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27
Db	6060	ATTATTGCTGGAATGAGCACAAGTTTAAAGAAATTCATTATTCAAGACTACCAATGAT	6119
Qy	28	PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
Db	6120	TTTGATGTATTAGATGTAATCAATTTATTTTGAAGATAAATACTATCTTGGTAAAAA	6179
Qy	48	ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys	67
Db	6180	TGCAAGGCAGTATTTTTCAATTCCTATTCTTTTTCGAACAATACTACACTTTAAAAACAT	6239
Qy	68	LeuIleGlnAenGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
Db	6240	TTAATCANAATCAAGAAATATGAGCCGAATTAATATATGTGTCTAATTACAAACCAAGCT	6299
Qy	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107
Db	6300	CATCTAGAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTCTGTGATGCTCATTTG	6359
Qy	108	GlyHisGluLeileLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle	127
Db	6360	GGATATGATTTTTTCAAAACAACCTAAAGATTTTAAATGCTTATTTTAAATTTACGAAATT	6419
Qy	128	TyrAsnArgGlnArgGluIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147
Db	6420	TATTTCAATCAAGAAATATACCTCAGGGGTCTATATGTGTCAGTAGCATTAGCCCTAGGA	6479

Percent Similarity:	69.62%	Conservative:	43
Best Local Similarity:	54.95%	Mismatches:	71
Query Match:	52.39%	Indels:	18
DB:	6	Gaps:	6
US-10-735-419-10 (1-294) x BD249790 (1-11474)			
QY	8	LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27
Db	6060	ATTATTGCTGGAAATGGACCAAGTTTAAAGAATTAATGATTTCAAGACTACCAATGAT	6119
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
Db	6120	TTTGATGATTTAGATTAATCAATTTTAAAGATAAATACTATCTTGGTAAAAA	6179
QY	48	ValLysTyrValPhePheAsnProPheValPheGluGlnTyrThrSerLysLys	67
Db	6180	TGCAAGCGCAGTATTTTACAATCTTATTTTGGAAACATACACATTTTAAAAACAT	6239
QY	68	LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
Db	6240	TTAATCCAAATCAAGATATGAGACCAACTAATTAATGTTCTTAATTACAACCAAGCT	6299
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107
Db	6300	CATCTAGAAAATGAAAAATTTGTAAAACTTTTACGATATTTTCTGATGCTCATTTG	6359
QY	108	GlyHisGluIleLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle	127
Db	6360	GGATATGATTTTTCACCAACTTAAAGATTTTAATGCTTATTTTAAATTTTCAAGAAAT	6419
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147
Db	6420	TATTTCAATCAAGAATTAACCTCAGGGTCTATATGTGTGAGTAGCCATAGCCCTAGA	6479
QY	148	TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla	167
Db	6480	TACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCACTCT--TATGCT	6536
QY	168	PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe	187
Db	6537	TTTGATCTATAACAAAAAATCTTTTAAATTTGGCTCTCTTAATTTTAAAAATGATTAATCA	6596
QY	188	LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg	207
Db	6597	CACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGATTTCTAGAAAAACT	6656
QY	208	TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu	223
Db	6657	TACAAAATAAAACTATATGCTTATGCTTAACTAAGCTTTTATAGCAAAATTTTATAGAACTA	6716
QY	224	AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp	243
Db	6717	GCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAATA-----ATAACTACACTAAAGAT	6773
LOCUS	AR271699	11474 bp	DNA
DEFINITION	Sequence 1 from patent US 6503744.		
ACCESSION	AR271699		
VERSION	AR271699.1		
RESULT 14			
LOCUS	AR271699	11474 bp	DNA
DEFINITION	Sequence 1 from patent US 6503744.		
ACCESSION	AR271699		
VERSION	AR271699.1		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 11474)		
AUTHORS	Gilbert, M. and Wakarchuk, W.W.		
TITLE	Campylobacter glycosyltransferases for biosynthesis of gangliosides and ganglioside mimics		
JOURNAL	Patent: US 6503744-A 1 07-JAN-2003;		
FEATURES	Location/Qualifiers		
source	1..11474		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ORIGIN			
Alignment Scores:	8.35e-56	Length:	11474
Pred. No.:	823.00	Matches:	161
Score:	69.62%	Conservative:	43
Percent Similarity:	54.95%	Mismatches:	71
Best Local Similarity:	52.39%	Indels:	18
Query Match:	6	Gaps:	6
DB:			
US-10-735-419-10 (1-294) x AR271699 (1-11474)			
QY	8	LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27
Db	6060	ATTATTGCTGGAAATGGACCAAGTTTAAAGAATTAATGATTTCAAGACTACCAATGAT	6119
QY	28	PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
Db	6120	TTTGATGATTTAGATTAATCAATTTTAAAGATAAATACTATCTTGGTAAAAA	6179
QY	48	ValLysTyrValPhePheAsnProPheValPheGluGlnTyrThrSerLysLys	67
Db	6180	TGCAAGCGCAGTATTTTACAATCTTATTTTGGAAACATACACATTTTAAAAACAT	6239
QY	68	LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
Db	6240	TTAATCCAAATCAAGATATGAGACCAACTAATTAATGTTCTTAATTACAACCAAGCT	6299
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107
Db	6300	CATCTAGAAAATGAAAAATTTGTAAAACTTTTACGATATTTTCTGATGCTCATTTG	6359
QY	108	GlyHisGluIleLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle	127
Db	6360	GGATATGATTTTTCACCAACTTAAAGATTTTAATGCTTATTTTAAATTTTCAAGAAAT	6419
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147
Db	6420	TATTTCAATCAAGAATTAACCTCAGGGTCTATATGTGTGAGTAGCCATAGCCCTAGA	6479
QY	148	TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla	167
Db	6480	TACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCACTCT--TATGCT	6536
QY	168	PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe	187
Db	6537	TTTGATCTATAACAAAAAATCTTTTAAATTTGGCTCTCTTAATTTTAAAAATGATTAATCA	6596
QY	188	LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg	207
Db	6597	CACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAACT	6656
QY	208	TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu	223
Db	6657	TACAAAATAAAACTATATGCTTATGCTTAACTAAGCTTTTATAGCAAAATTTTATAGAACTA	6716
QY	224	AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp	243
Db	6717	GCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAATA-----ATAACTACACTAAAGAT	6773
QY	244	IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly-----	257
Db	6774	ATACTATACCTTCTAGTGAGGCT-----TATGGAAAAATTTTCAAAAAATATAT	6821
QY	258	-----LysLysSerArgLeuLysGluAsnLeuHieTyrLysLeuIleLysAspLeuIle	275
Db	6822	AAATTTTAAAAAATAAAATAAAGAAATATATTTATACAGTTGATAAAGATCTATTAA	6881
QY	276	ArgLeuProSerAspIleLysHisTyrLeuLysGluLys	288
Db	6882	AGATTACCTAGTATAAAGCATTAATTTCAAGGAAAA	6920

```
QY 244 IleuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 6774 ATACTCATACCTTCTAGTGGGCT-----TATGAAATTTTCAAAAAATATT 6821
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuileLysAspLeuile 275
Db 6822 AATTTTAAAAAATAAAATTAAGAAATATTATTACAGTTGATAAAGATCTATT 6881
QY 276 ArgLeuProSerAspLysHisTyrLeuLysGluLys 288
Db 6882 AGATTACCTAGTATATAAGCATTATTTCAGAGGAAAA 6920

RESULT 15
AR481781
LOCUS AR481781 11474 bp DNA linear PAT 14-MAY-2004
DEFINITION Sequence 1 from patent US 6699705.
ACCESSION AR481781
VERSION AR481781.1 GI:47243416
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 11474)
AUTHORS Gilbert M. and Wakarchuk W.W.
TITLE Campylobacter glycosyltransferases for biosynthesis of gangliosides
and ganglioside mimics
JOURNAL Patent: US 6699705-A 1 02-MAR-2004;
FEATURES
source
1. 11474
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 8.35e-56 Length: 11474
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 6 Gaps: 6

US-10-735-419-10 (1-294) x AR481781 (1-11474)

QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 6060 ATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTATTCAGAGACTTACCAAAATGAT 6119
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 6120 TTTGATGATTATAGATGTAATCAATTTTATTTGAAGATATAATCTATCTTGGTAAAAA 6179
QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 6180 TGCAAGCGCAGTATTTTACATCTCTATCTTTTGAACAATCTACACTTTTAAACAT 6239
QY 68 LeuileGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 6240 TTAATCCAAATCAAGAAATATGAGACCGAAGTAATTTATGTTCTTAATTACACCAAGCT 6299
QY 88 TyriLeaspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 6300 CATCTAGAAATGAAATTTTGAAGAACTTTTACGATTATTTCTGATGCTCATTTG 6359
QY 108 GlyHisGluileLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluile 127
Db 6360 GGATATGATTTTTCACCAACTTAAGATTTTAAATGCTTATTTTAAATTCACGAAAT 6419
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 6420 TATTTCAATCAAGAATTACCTCAGGGCTCTATATGTGTGCTAGTAGCCCTAGGA 6479
QY 148 TyriLysSerileTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyrAla 167
```

```
Db 6480 TACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 6536
QY 168 PheAspAsnAsnLysLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 6537 TTTGATACTAAACAAAAAATCTTTTAAAAATGGCTCCTAAATTTTAAAAATGATAATTCA 6596
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 6597 CACTATATCGGACATAGTAAAAATACAGATATATAAAGCTTTTAGAATTTTCTAGAAAAAACT 6656
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 6657 TACAAATTAACACTATATTGCTTATGTCCTACAGCTTTTACCAATTTTATAGAACTA 6716
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 6717 GCGCCAAATTTAAATTTCAAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 6773
QY 244 IleLeuileProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 6774 ATACTCATACCTTCTAGTGGGCT-----TATGGAAAAATTTTCAAAAAATATT 6821
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLysLeuileLysAspLeuile 275
Db 6822 AATTTTAAAAAATAAAATTAAGAAATATTATTATCAAGTTGATAAAGATCTATT 6881
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 6882 AGATTACCTAGTATATAAAGCATTATTTCAGAGGAAAA 6920
```

Search completed: June 14, 2005, 20:11:30
Job time : 5010 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 16:43:29 ; Search time 612 Seconds
(without alignments)
2843.797 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNINALVCGNGPSLKNIID.....IRLPDIKHYLKEKYANKNR 294

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlh
-Q=/cgn2_1/USPTO.spool/US10735419/runat_13062005_165724_12636/app_query.fasta_1.455
-DB=N Geneseq_16Dec04 -QFMT=fastap -SUFFIX=rng -MINMATCH=0.1 -LOOPEGL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10735419 @CGN 1.1 644 @runat_13062005_165724_12636 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	823	52.4	876	3	Aaa53721
2	823	52.4	876	3	Abt13666 C. jejuni
3	823	52.4	11474	3	Aaa53720 LPS core
4	823	52.4	11474	8	Abt13665 Campyloba
5	817	52.0	876	3	Aaa53725 Campyloba

6	817	52.0	876	8	Abt13668 C. jejuni
7	813	51.8	876	3	Aaa53724 Campyloba
8	813	51.8	876	3	Abt13667 C. jejuni
9	804	51.2	873	8	Abt13670 C. jejuni
10	804	51.2	873	8	Abt13671 C. jejuni
11	804	51.2	876	3	Aaa53726 Campyloba
12	804	51.2	876	3	Abt13669 C. jejuni
13	680.5	43.3	1293	2	Aaz25693 Campyloba
14	673.5	42.9	909	8	Acc71694 Haemophil
15	663.5	42.2	19196	13	Continuation (4 of
16	662.5	42.2	110000	2	AAT42063_03
17	653.5	41.6	13379	13	ADT05530 Haemophil
18	653.5	41.6	117576	13	ADT05736 Haemophil
19	653.5	41.6	34980	13	ADT05648 Haemophil
20	609.5	38.8	4277	13	ADT05421 Haemophil
21	128.5	8.2	6042	3	AAA70199 Plasmodi
22	127.5	8.1	50000	6	ABL56202 AmEPV gen
23	124.5	7.9	4197	3	AAA70227 Plasmodi
24	123	7.8	2286	3	AAA69014 Bacteriop
25	123	7.8	1668	3	AAA69013 Bacteriop
26	122	7.8	50000	6	ABL56201 AmEPV gen
27	121.5	7.7	2825	2	AAQ14833 HincII re
28	121.5	7.7	3549	3	AAA70223 Plasmodi
29	119	7.6	675	8	ACA28886 Prokaryot
30	119	7.6	110000	2	AXX20248_05
31	119	7.6	116277	2	AXX20249 Borrelia
32	118.5	7.5	50000	6	ABL55643 AmEPV gen
33	116.5	7.4	1062	2	AAK61556 B. burgdo
34	116.5	7.4	1122	2	AAK61555 B. burgdo
35	116.5	7.4	5454	3	AAA70236 Plasmodi
36	116.5	7.4	110000	2	AXX20248_00
37	116.5	7.4	110000	2	AXX20248_01
38	116	7.4	2181	6	ABL56216 AmEPV NTP
39	116	7.4	50000	6	ABL55644 AmEPV gen
40	115	7.3	885	6	ABL55636 AmEPV fir
41	115	7.3	3366	3	AAA70175 Plasmodi
42	115	7.3	50000	6	ABL55644 AmEPV gen
43	114.5	7.3	3872	2	AAX23700 P. falcip
44	114.5	7.3	3872	3	AAX82995 P. falcip
45	114	7.3	2616	8	ACA27544 Prokaryot

ALIGNMENTS

RESULT 1
AAA53721 standard; DNA; 876 BP.

XX AAA53721;

XX 15-SEP-2003 (revised)
DT 22-DEC-2000 (first entry)

XX Campylobacter jejuni OH4384 CstII sialyltransferase coding sequence.

XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAC transferase;
KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
KW immunity; immunogen; ganglioside; ds.

XX Campylobacter jejuni; OH4384.

XX Key Location/Qualifiers

FT CDS 1..876
FT /tag= a
FT /product= "CstII sialyltransferase"

XX WO200046379-A1.

PD 10-AUG-2000.

XX 01-FEB-2000; 2000WO-CA000086.

XX New glycosyltransferases from *Campylobacter*, useful for synthesizing
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 PT mechanisms of organisms that synthesize ganglioside mimetics.
 XX
 XX
 XX Claim 8; Page 96; 107pp; English.

CC The invention comprises the amino acid and coding sequences of
 CC *Campylobacter jejuni* proteins. The C. *jejuni* proteins of the invention
 CC may be either an: acetyltransferase; glycosyltransferase; GalNAc (N-
 CC Acetyl-galactosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthetase; acetyltransferase. The C. *jejuni* DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of *Campylobacter* enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. *jejuni* oligosaccharides of the invention may be used as
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
 CC metastasis). The present DNA sequence represents a *Campylobacter jejuni*
 CC gene of the invention

XX SQ Sequence 876 BP; 359 A; 110 C; 103 G; 304 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.2e-73 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Best Local Similarity: 54.95% Mismatches: 71
 Query Match: 52.39% Indels: 18
 DB: 8 Gaps: 6

US-10-735-419-10 (1-294) x ABT13666 (1-876)

Qy 8 LeuValCysGlyAenGlyProSerLeuLysAsnLeuAspTyrLysArgLeuProLysGln 27
 Db : : : : :
 Qy 13 ATTATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTCACAGACTACCAATGAT 72
 Db : : : : :
 Qy 28 PhaAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db : : : : :
 Qy 73 TTTGATGATTATTAGATGAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAAA 132
 Db : : : : :
 Qy 48 ValLysTyrValPhePheAenProPheValPhePheGluGlnTyrTyrSerLysLys 67
 Db : : : : :
 Qy 133 TGCAGGCGAGTATTTTCAATCTCTATTTTGAAGATAAATACTATCTTGGTAAAAAA 192
 Db : : : : :
 Qy 68 LeuLeuGlnAenGluTyrAsnLysGluAsnLeuValCysSerThrLysAsnLeuGlu 87
 Db : : : : :
 Qy 193 TTAATCCAAATCAAGATATGAGCCGAATAATTATGTGTTCTAATTACCAAGCT 252
 Db : : : : :
 Qy 88 TyrLysAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db : : : : :
 Qy 253 CATCTAGAAATGAAATTTTGTAAAACTTTTACGATTTATTTCTGATGCTCATTTG 312
 Db : : : : :
 Qy 108 GlyHisGluLeuLeuLysLysLeuLysAspPhePheAlaTyrLysTyrAsnGluLeu 127
 Db : : : : :
 Qy 313 GGATATGATTTTTCAAACCACTTAAAGATTTTAAATGCTTATTTTAAATTTTCGAAT 372
 Db : : : : :
 Qy 128 TyrAsnArgGlnArgLysThrSerGlyValTyrMetCysAlaThrAlaAlaLeuGly 147
 Db : : : : :
 Qy 373 TATTTCAATCAAGATATCTCAGGGGTTTATATGTGTCAGTAGCATAGCCCTAGGA 432
 Db : : : : :
 Qy 148 TyrLysSerLysTyrLysSerGlyLysAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
 Db : : : : :
 Qy 433 TACAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGTCACTCT--TATGCT 489
 Db : : : : :
 Qy 168 PheAspAsnAsnLysLysAsnLeuLysAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 Db : : : : :
 Qy 490 TTTGATACTAAACAAATAAATCTTTTAAATTTGGCTCTCTCTAATTTTAAATAATGATAATCA 549
 Db : : : : :
 Qy 188 LysPheLysAsnHisSerMetAlaCysAspLysGlnAlaLeuAspTyrLeuMetLysArg 207
 Db : : : : :
 Qy 550 CACTATATCGACATATGATAAATAACAGATATATAAGCTTTAGATTTTCTAGAAAAACT 609

Qy 208 TyrAspValAsnLysTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 Db : : : : :
 Qy 610 TACAAATAAACTATATTGCTTATGCTCTTAAACAGTCTTTTACAAATTTTATAGAACTA 669
 Db : : : : :
 Qy 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrLysTyrLysAsnAsp 243
 Db : : : : :
 Qy 670 GCGCCAAATTAATTCAAATTTTATCATACACAGAAAA---AATAACTACACTAAAGAT 726
 Db : : : : :
 Qy 244 IleLeuLeuProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 Db : : : : :
 Qy 727 ATACTCATACCTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
 Db : : : : :
 Qy 258 -----LysLysSerArgLysGluAsnLeuHisTyrLysLysLeuLysAspLeuLeu 275
 Db : : : : :
 Qy 775 AATTTTAAAAAATAAATAATTTAAAGAAATATATTATTATACAAAGTTGATAAAAGATCTATTA 834
 Db : : : : :
 Qy 276 ArgLeuProSerAspIleLysHisTyrLysLysLysGluLys 288
 Db : : : : :
 Qy 835 AGATTACCTAGTGATATTAAGCATTTATTTCAAGGAAAA 873

RESULT 3
 AAA53720
 ID AAA53720 standard; DNA; 11474 BP.
 XX
 AC AAA53720;
 XX
 DT 15-SEP-2003 (revised)
 DT 22-DEC-2000 (first entry)
 XX
 DE LPS core biosynthesis locus.
 XX
 KW Biosynthetic locus; biosynthesis; lipid A biosynthesis;
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
 KW immunity; immunogen; ganglioside; ds.
 XX
 OS *Campylobacter jejuni*; OH4384.
 XX
 Key Location/Qualifiers
 CDS 350..1237
 FT /tag= a
 FT /product= "acyltransferase"
 FT /note= "Open reading frame 2b"
 FT 1234..2490
 CDS
 FT /tag= b
 FT /product= "glycosyl transferase"
 FT /note= "Open reading frame 3a"
 FT 2786..3355
 CDS
 FT /tag= c
 FT /product= "Glycosyl transferase"
 FT /note= "Open reading frame 4b"
 FT 4025..5068
 CDS
 FT /tag= d
 FT /product= "Beta 1,4-GalNAc transferase"
 FT /note= "Open reading frame 5a"
 FT 5054..5959
 CDS
 FT /tag= e
 FT /product= "Beta 1,2- galactosyltransferase"
 FT /note= "Open reading frame 6a"
 FT 6048..6923
 CDS
 FT /tag= f
 FT /product= "CstII sialyltransferase"
 FT /note= "Open reading frame 7a"
 FT 6924..7964
 CDS
 FT /tag= g
 FT /product= "Sialic acid synthase"
 FT /note= "Open reading frame 8a"
 FT 8020..9079
 CDS
 FT /tag= h
 FT /product= "Sialic acid biosynthetic enzyme"
 FT /note= "Open reading frame 9a"

DR P-PSDB; AAY97211.

Novel glycosyltransferase polypeptides and polynucleotides useful for biosynthesis of ganglioside and ganglioside mimics, as diagnostic reagents and as immunogen for producing antibodies.

PS Claim 6; Page 94: 120pp: English.

A reaction mixture for the synthesis of a sialylated oligosaccharide is useful for synthesising sialylated oligosaccharides such as ganglioside, lyso-ganglioside or their mimics. Glycosyltransferases are useful for chemo-enzymatic synthesis of oligosaccharides, including gangliosides and other oligosaccharides that have biological activity. The enzymes and nucleic acids that encode them are useful for studies of the pathogenesis and mechanisms of organisms that synthesise ganglioside mimics, such as *C. jejuni* and the nucleic acids are used as probes to study expression of genes involved in ganglioside mimetic synthesis. Antibodies raised against the glycosyltransferases are also useful for analysing the expression patterns of these genes involved in pathogenesis. The nucleic acids are also useful for designing antisense oligonucleotides for inhibiting expression of the Campylobacter enzymes that are involved in the biosynthesis of ganglioside mimics that can mask the pathogens from the host's immune system. The oligosaccharides are useful as diagnosing reagents or as therapeutics and as immunogens for producing antibodies. Bacterial glycosyltransferase can be used to catalyse the formation of oligosaccharides that are identical to the corresponding mammalian structures and are easier and less expensive to produce in large quantity, compared to the mammalian glycosyltransferase. The bacterial origin of the enzymes facilitates expression of large quantities of the enzymes using relatively inexpensive prokaryotic expression systems. (Updated on 15-SEP-2003 to standardise OS field)

SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8,83e-73	Length:	876
Score:	817.00	Matches:	159
Percent Similarity:	69.62%	Conservative:	45
Best Local Similarity:	54.27%	Mismatches:	71
Query Match:	52.01%	Indels:	18
DB:	3	Gaps:	6

US-10-735-419-10 (1-294) x AAA53725 (1-876)

Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArqLeuProLysGln 27

Db
13 ATTATGCTGGAAATGGACCAAGCTTTAAAGAAATTGATTATTTCAAGACTACCAAATGAT 72

Qy 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47

Db 73 TTTGATGTA TTTAGATGCAATCAATTTTATTTTGAAGATAAAATACTATCTTTGGTAAAAA 132

Qy 48 VallLysTyrValPhePheAsnPropheValPhePheGluGlnTyrTyrThrSerLysLys 67

Db 133 TGCAAAGCAGTATTTTACAATCCTAGTCTTTTTTTTGAAACAATACTACACTTTTAAACAAT 192

Qy 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87

D**b** 193 TTAATCCAAATCAAGAATATGAGACCGAACTAATCATGTGTTCTAATTTTAACCAAGCT 252

Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107

D**b** 253 CATCTAGAAAAATCAAAAATTTTGTAAAAAATTTTACGATATATTTCTGATGCTCATTTG 312

Qy 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127

Db 313 GGATATGATTTTTCAAACAACCTTAAAGAAATCAATGCTTATTTTAAATTCACGAAAT 372

Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147

Db 373 TATTTCATCAAGAATTACCTCAGGGGCTATATGTGCACAGTAGCCATAGCCCTAGGA 432

Qy 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167

CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-Acetylglucosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP) sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein sequences of the invention are useful for ganglioside synthesis, studying ganglioside mimetics, and for designing oligonucleotides to inhibit expression of Campylobacter enzymes involved in the biosynthesis of ganglioside mimetics that can mask the pathogen's from the host's immune system. The C. jejuni oligosaccharides of the invention may be used as diagnostic reagents (e.g. to locate areas of inflammation or tumour metastasis). The present DNA sequence represents a Campylobacter jejuni CC gene of the invention

XX SQ Sequence 876 BP; 361 A; 114 C; 102 G; 299 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,83e-73 Length: 876
 Score: 817.00 Matches: 159
 Percent Similarity: 69.62% Conservative: 45
 Best Local Similarity: 54.27% Mismatches: 71
 Query Match: 52.01% Indels: 18
 DB: Gaps: 6

US-10-735-419-10 (1-294) x ABT13668 (1-876)

QY 8 LeuValCysGlyAenGlyProSerLeuLysAenIleAspTyrLysArgLeuProLysGln 27
 DB 13 ATTATTGCTGGAATGACCAAGTCTTAAAGAAATGATTATTCAGAGACTACCAAAATGAT 72
 QY 28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 DB 73 TTTGATGATTTAGATGCAATCAATTTTATTTGAAGATATAATACATCTTGTGTAAGAAA 132
 QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
 DB 133 TGCAGAGCAGTATTTACATCTAGTCTTTTTTTTGAACAATACATCTTTAAACAT 192
 QY 68 LeuIleGlnAenGluTyrAenIleGluAenIleValCysSerThrIleAenLeuGlu 87
 DB 193 TTAAATCAAAATCAAGAAATATGAGACCGCACTAATCATGTGTCTTAATTTTACCAAGCT 252
 QY 88 TyrIleAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
 DB 253 CATCTAGAAATCAAAATTTTGTAAGAACTTTTACAGTATTTTCTGATGCTCATTTG 312
 QY 108 GlyHisGluIleLysLysLeuLysAspPheAlaTyrIleLysTyrAenGluIle 127
 DB 313 GGATATGATTTTTTCAACAACCTTAAAGAAATTCATTTTAAATTTTCAGAAAT 372
 QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValLalaLeuGly 147
 DB 373 TATTTCAATCAAGAATTTACCTCAGGGGTCTATATGTGCACAGTAGCCATAGCCCTAGGA 432
 QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAenAenLeuTyrAla 167
 DB 433 TACAAAGAAATTTATCTTCGGGAATTTGATTTTTTCAAAATGATCATCT---TATGCT 489
 QY 168 PheAspAenAsnLysLysAenLeuLysAenLysCysThrGlyPheLysAenGlnLysPhe 187
 DB 490 TTTGATACCAACAACAAATAATCTTTTAAATTTGGCTCCTTAATTTTAAATGATAATTC 549
 QY 188 LysPheIleAenHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 DB 550 CACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAA 609
 QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 DB 610 TACGAATAAAGTATATTTGTTTATGTCTTAAACAGTCTTTTACCAAAATTTATAGAACTA 669
 QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysTyrIleAenAsp 243
 DB 670 GCGCCAAATTTAAATTCNAATTTTATCATACAGAAAAA---NATAACTATATAAGAT 726

QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 DB 727 ATACTCATACCTCTAGTCAGGCT-----TATGGAATAATTTACAAAAATATT 774
 QY 258 -----LysLysSerArgLeuLysGluAenLeuHisTyrLysLeuIleLysAspLeuIle 275
 DB 775 AATTTTAAATAATAAATAATTAAGAAATATATTATTACAGTTGATAAAGATCTATT 834
 QY 276 ArgLeuProSerAspIleLysHisTyrLysLysGluLys 288
 DB 835 AGATTACCTAGTATATAAGCATTTATTTCAAAGGAAAA 873
 RESULT 7
 ID AAAS3724
 XX AAAS3724 standard; DNA; 876 BP.
 AC AAAS3724;
 DT 15-SEP-2003 (revised)
 DT 22-DEC-2000 (first entry)
 DE Campylobacter jejuni O:10 serotype alpha-2,3-sialyltransferase.
 XX Biosynthetic locus; biosynthesis; lipid A biosynthesis;
 KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;
 KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;
 KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;
 KW immunity; immunogen; ganglioside; ds.
 XX Campylobacter jejuni; O:10 serotype.
 XX
 FH Location/Qualifiers
 FT 1..876
 FT /*tag= a
 FT /product= "alpha-2,3-sialyltransferase"
 PN WO200046379-A1.
 PD 10-AUG-2000.
 XX
 PF 01-FEB-2000; 2000WO-CA000086.
 PR 01-FEB-1999; 99US-0118213P.
 PR 31-JAN-2000; 2000US-00495406.
 XX (CANA) NAT RES COUNCIL CANADA.
 XX Gilbert M, Wakarchuk WW;
 XX WPI; 2000-524418/47.
 DR P-PSDB; AAY97210.
 XX
 PT Novel glycosyltransferase polypeptides and polynucleotides useful for
 PT biosynthesis of ganglioside and ganglioside mimics, as diagnostic
 PT reagents and as immunogen for producing antibodies.
 XX
 PS Claim 6; Page 92; 120pp; English.
 XX
 CC A reaction mixture for the synthesis of a sialylated oligosaccharide is
 CC useful for synthesizing sialylated oligosaccharide such as ganglioside,
 CC lysoganglioside or their mimics. Glycosyltransferases are useful for
 CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
 CC other oligosaccharides that have biological activity. The enzymes and
 CC nucleic acids that encode them are useful for studies of the pathogenesis
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
 CC jejuni and the nucleic acids are used as probes to study expression of
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analyzing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the Campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing

CC reagents or as therapeutics and as immunogens for producing antibodies.
 CC Bacterial glycosyltransferase can be used to catalyze the formation of
 CC oligosaccharides that are identical to the corresponding mammalian
 CC structures and are easier and less expensive to produce in large
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial
 CC origin of the enzymes facilitates expression of large quantities of the
 CC enzymes using relatively inexpensive prokaryotic expression systems.
 CC (Updated on 15-SEP-2003 to standardise OS field)
 XX

SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,23e-72 Length: 876
 Score: 813.00 Matches: 159
 Percent Similarity: 69.97% Conservative: 46
 Best Local Similarity: 54.27% Mismatches: 70
 Query Match: 51.75% Indels: 18
 DB: Gaps: 6

US-10-735-419-10 (1-294) x AAA53724 (1-876)

QY	8	LeuValCysGlyAenGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln	27
DB	13	ATTATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAGAGGCTACCAATGAT	72
QY	28	PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp	47
DB	73	TTTGATGATTATAGATGCAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA	132
QY	48	ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrThrSerLysLys	67
DB	133	TTCAAAGCAGTATTTTACAATCTGGTCTTTTTTTTGAACAATACTACACTTTAAACAT	192
QY	68	LeuIleGlnGluGlnTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu	87
DB	193	TTAATCAAATCAAGATATGAGCCGAATTAATTTATGTGTCTAATTACAAACCAAGCT	252
QY	88	TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu	107
DB	253	CATCTAGAAATGAAATTTGTAAAACTTTTACGATTTTCTGATGCTCATTTG	312
QY	108	GlyHisGluIleLysLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle	127
DB	313	GGATATGATTTTTTAAACAACCTAAAGAAATTAATGCTTATTTTAAATTTCAACGAAAT	372
QY	128	TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly	147
DB	373	TATCTCAATCAAGAATTAACCTCAGGAGCTATATGTGTGCTAGCTATAGCCCTAGGA	432
QY	148	TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla	167
DB	433	TACAAAGAAATTTATCTTCTGGAATTTATTTATCAAAATGGGTCATCT--TATGCT	489
QY	168	PheAspAsnAsnLysLysAsnLeuLysLysCysThrGlyPheLysAsnGlnLysPhe	187
DB	490	TTTGATACCAACAAGAAATCTTTTAAAACTGGCTCTCGAATTTAAAAATGATCGCTCA	549
QY	188	LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg	207
DB	550	CACATATCGGACATAGTAAAAATACAGATATAAAGCTTTAGATTCTCAGAAAAACT	609
QY	208	TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu	223
DB	610	TACAAATAAAACTATATTGCTTATGCTCTACAGCTCTTTAGCAAAATTTTATAGACTA	669
QY	224	AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp	243
DB	670	GGCCCAATTTAAATTTCAAAATTTTATCATACAAGAAAAA---AATAACTACACTAAAGAT	726
QY	244	IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly-----	257
DB	727	ATACTCATACCTTCTAGTGAGGCT-----TATGGAAATTTTCAAAAAATATT	774

QY	258	-----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle	275
DB	775	ATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA	834
QY	276	ArgLeuProSerAspIleLysHisTyrLeuLysGluLys	288
DB	835	AGATTACCTAGTGATATAAAGCATTATTTTCAAAAGGAAAA	873

RESULT 8

ABTI13667
 ID ABTI13667 standard; DNA; 876 BP.

AC ABTI13667;

XX 07-FEB-2003 (first entry)

DE C. jejuni bifunctional sialtransferase cstII coding sequence #2.

KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
 KW GalNAc transferase; N-Acetylglucosamine transferase;
 KW galactosyltransferase; sialyltransferase; sialic acid synthase;
 KW cytidine 5'-monophosphate sialic acid synthetase;
 KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
 KW ganglioside mimetics; inflammation; tumour metastasis.

OS Campylobacter jejuni.

XX WO200274942-A2.

XX 26-SEP-2002.

PF 22-FEB-2002; 2002WO-CA000229.

PR 21-MAR-2001; 2001US-00816028.

XX (CANA) NAT RES COUNCIL CANADA.

PI Gilbert M, Wakarchuk WW;

DR WPI; 2003-040554/03.

DR P-PSDB; ABU18480.

XX New glycosyltransferases from Campylobacter, useful for synthesizing
 PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
 PT mechanisms of organisms that synthesize ganglioside mimetics.

XX Claim 8; Page 96-97; 107pp; English.

CC The invention comprises the amino acid and coding sequences of
 CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
 CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
 CC Acetylglucosamine) transferase; galactosyltransferase;
 CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
 CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
 CC sequences of the invention are useful for ganglioside synthesis, studying
 CC ganglioside mimetics, and for designing oligonucleotides to inhibit
 CC expression of Campylobacter enzymes involved in the biosynthesis of
 CC ganglioside mimetics that can mask the pathogen's from the host's immune
 CC system. The C. jejuni oligosaccharides of the invention may be used as
 CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
 CC metastasis). The present DNA sequence represents a Campylobacter jejuni
 CC gene of the invention

SQ Sequence 876 BP; 356 A; 115 C; 105 G; 300 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,23e-72 Length: 876
 Score: 813.00 Matches: 159
 Percent Similarity: 69.97% Conservative: 46
 Best Local Similarity: 54.27% Mismatches: 70
 Query Match: 51.75% Indels: 18
 DB: Gaps: 6


```

US-10-735-419-10 (1-294) x ABT13667 (1-876)
Qy      8 LeuValCysGlyAenGlyProSerLeuLysAenLysAspTyrLysArgLeuProLysGln 27
Db      13 ATTATTGCTGGAATGACCAAGTTTAAAGAAATTCATTATTCAAGGCTACCAAAATGAT 72
Qy      28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGATTAGATGCAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAAA 132
Qy      48 ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db      133 TTCAAACAGTATTTTACAACTCTGGTCTTTTTCGACCAATACACTTTTAAACAT 192
Qy      68 LeuLeuGlnAenGluLysAenLysGluAenLysValCysSerThrLysAenLeuGlu 87
Db      193 TTAATCCAAAATCAAGAATATGAGACCGCAATTAATATGTTGTTCTTAATTACAAACCAAGCT 252
Qy      88 TyrIleAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAATGAAATTTTGTAAACCTTTTACGATTATTTCTGATGCTCATTTG 312
Qy      108 GlyHisGluLeuLysLysLeuLysAenPhePheAlaTyrLysTyrAenGluLeu 127
Db      313 GGATATGATTTTAAACAACTTAAAGAAATTTAATGCTTTATTTTAAATTTACGAAAT 372
Qy      128 TyrAenArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATCTCAATCAAGAAATTAAGTCTGAGGAGTCTATATGTTGTCAGTAGCTATAGCCCTAGGA 432
Qy      148 TyrIleSerIleTyrLysGlyLeuAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db      433 TACAAGAAATTTATCTTTCTGGAATGATTTTATCAAAATGGGTCACTCT---TATGCT 489
Qy      168 PheAspAenAenLysLysAenLeuLysCysThrGlyPheLysAenGlnLysPhe 187
Db      490 TTTGATACCAACCAAGAAATCTTTAAACTGGCTCTGATTTTAAAAATGATCGCTCA 549
Qy      188 LysPheIleAenHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAenLysSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAATAAATCAATATATGCTTATGCTTATGCTTCAACAGTCTTTTAGCAAAATTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTCAAATTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuLeuProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTCTAGTAGGCT-----TATGCAAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAenLeuHisTyrLysLysLeuLysAspLeuLeu 275
Db      775 AATTTTAAAAAATAAATAAATAAAGAAATATTATTACAAAGTGTATAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLysLysLysGluLys 288
Db      835 AGATTACTAGTATAGATTAAGCATTTATTTCAAGGAAAA 873

RESULT 9
ID      ABT13670
XX      XX
AC      ABT13670;
XX      XX
DT      07-FEB-2003 (first entry)
XX      XX
DE      C. jejuni bifunctional sialtransferase cstII coding sequence #5.
XX      XX
KW      Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;

```

```

KW      GalNAC transferase; N-Acetylgalactosamine transferase;
KW      galactosyltransferase; sialyltransferase; sialic acid synthase;
KW      cytidine 5'-monophosphate sialic acid synthetase;
KW      CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW      ganglioside mimetics; inflammation; tumour metastasis.
XX      Campylobacter jejuni.
XX      WO200274942-A2.
XX      26-SBP-2002.
XX      22-FEB-2002; 2002WO-CA000229.
XX      21-MAR-2001; 2001US-00816028.
XX      (CANA ) NAT RES COUNCIL CANADA.
XX      Gilbert M, Wakarchuk WW;
XX      WPI; 2003-040554/03.
XX      P-PSDB; ABJ18484.
XX      New glycosyltransferases from Campylobacter, useful for synthesizing
XX      gangliosides and ganglioside mimetics, and in studying the pathogenesis
XX      mechanisms of organisms that synthesize ganglioside mimetics.
XX      Disclosure; Page 98-99; 107pp; English.
XX      The invention comprises the amino acid and coding sequences of
XX      Campylobacter jejuni proteins. The C. jejuni proteins of the invention
XX      may be either an: acyltransferase; glycosyltransferase; GalNAC (N-
XX      Acetylgalactosamine) transferase; galactosyltransferase;
XX      sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
XX      sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
XX      sequences of the invention are useful for ganglioside synthesis, studying
XX      ganglioside mimetics, and for designing oligonucleotides to inhibit
XX      expression of Campylobacter enzymes involved in the biosynthesis of
XX      ganglioside mimetics that can mask the pathogen's from the host's immune
XX      system. The C. jejuni oligosaccharides of the invention may be used as
XX      diagnostic reagents (e.g. to locate areas of inflammation or tumour
XX      metastasis). The present DNA sequence represents a Campylobacter jejuni
XX      gene of the invention
XX      SQ      Sequence 873 BP; 350 A; 118 C; 110 G; 295 T; 0 U; 0 Other;

```

Alignment Scores:

Pred. No.:	1-79e-71	Length:	873
Score:	804.00	Matches:	158
Percent Similarity:	69.28%	Conservative:	45
Best Local Similarity:	53.92%	Mismatches:	72
Query Match:	51.18%	Indels:	18
DB:	8	Gaps:	6

US-10-735-419-10 (1-294) x ABT13670 (1-873)

```

Qy      8 LeuValCysGlyAenGlyProSerLeuLysAenLysAspTyrLysArgLeuProLysGln 27
Db      13 ATTATTGCTGGAATGACCAAGTTTAAAGAAATTCATTATTCAAGGCTACCAAAATGAT 72
Qy      28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGATTAGATGCAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAAA 132
Qy      48 ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db      133 TCGCAAGCAGTGTTTTACACCCCTGGTCTTCTTTGACCAATACACTTTTAAACAT 192
Qy      68 LeuLeuGlnAenGluTyrAenLysGluAenLysValCysSerThrLysAenLeuGlu 87
Db      193 TTAATCCAAAATCAAGAATATGAGACCGCAATTAATATGTTGTTCTTAATTACAAACCAAGCT 252
Qy      88 TyrIleAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107

```

```

Db 253 CATCTAGAAATGAAATTTTGTAAACACTTTTACGATATTATTTCTGATGCTCATTTG 312
Qy GlyHisGluIleLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGATATGATTTTAAACAACTTAAAGAAATTTAAATGCTTATTTTAAATTTTCAGAAAT 372
Qy TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAAATTAATCTAGGGGTCTATATGTGTGATGATGAGCCATAGGCTAGGA 432
Qy TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
Qy PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTCATTTTAAAAATGATCGCTCA 549
Qy LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAATAAACT 609
Qy TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAACTATATATGCTTATGTCCTACAGTCTTTTAGCAAAATTTTATAGAACTA 669
Qy AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 GCGCCAAATTTAAATTCMAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
Qy IleLeuIleProAspLysTyrAlaGlnGluArgTyrGly----- 257
Db 727 ATACTCATACCTCTAGTGGCT-----TATGCAAAATTTTCAAAATAATTT 774
Qy 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db 775 AATTTTAAAAAATAAAATTAAGAAAAATGTTTATTACAAGTTGATAAAAGATCTATTA 834
Qy ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTATATAAGCATTTATTTCAAGGAAAA 873

RESULT 10
ABT13671
ID ABT13671 standard; DNA; 873 BP.
XX
AC ABT13671;
XX
DT 07-FEB-2003 (first entry)
DE
DE C. jejuni bifunctional sialtransferase cstII coding sequence #6.
XX
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
OS Campylobacter jejuni.
XX
PN WO200274942-A2.
XX
PD 26-SEP-2002.
XX
PD 22-FEB-2002; 2002WO-CA000229.
XX
PF
PR 21-MAR-2001; 2001US-00816028.
XX
PA (CANVA ) NAT RES COUNCIL CANADA.
XX
XX Gilbert M, Wakarchuk WW;

```

```

XX
DR WPI; 2003-040554/03.
DR P-PSDB; ABU18485.
XX
PT New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS Disclosure; Page 99; 107pp; English.
XX
CC The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present DNA sequence represents a Campylobacter jejuni
CC gene of the invention
XX
SQ Sequence 873 BP; 353 A; 117 C; 107 G; 296 T; 0 U; 0 Other;

```

```

Alignment Scores:
Pred. No.: 1.79e-71 Length: 873
Score: 804.00 Matches: 158
Percent Similarity: 69.28% Conservative: 45
Best Local Similarity: 53.92% Mismatches: 72
Query Match: 51.18% Indels: 18
DB: 8 Gaps: 6

```

US-10-735-419-10 (1-294) x ABT13671 (1-873)

```

Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATTATTTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGGCTACCAATGAT 72
Qy 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGATTTAGATGTAATCAATTTATTTTGAAGATAAATACTATCTTGGTAAAAA 132
Qy 48 ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCAGAAACAGTGTTTTACACCCCTAATTTCTTCTTTGAGCAATACTACACTTTAAAAACAT 192
Qy 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAATCCAAATCAAGAAATATGAGCCGAACTAAATATATGTGTCTTAATTTACAAACCAAGCT 252
Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAATGAAATTTTGTAAACACTTTTACGATTTATTTCTGATGCTCATTTG 312
Qy 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db 313 CGATATGATTTTAAACCAACTTAAAGAAATTTAATGCTTATTTTAAATTTTACCAGAAAT 372
Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAAATTAATCTAGGGGTCTATATGTGTGATGATGAGCCATAGGCTAGGA 432
Qy 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
Qy 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTGATACCAACCAAGAAATCTTTTAAAACTAGCCCTCATTTTAAAAATGATCGCTCA 549

```

QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 Db 550 CACTATATCGCATAGTAAAAATACAGATATAAAGCTTTGAAATTTCTAGAAAAAACT 609
 QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 Db 610 TACAAAAATAAACATATATTGCTTATGTCCTTAATAGTCTTTTACCAAAATTTTATAGAACTA 669
 QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
 Db 670 GCGCCAAATTAATTAATTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
 QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 Db 727 ATATCATACCTTCTAGTAGGCT-----TATGGAAAAATTTCAAAAAATATT 774
 QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
 Db 775 AATTTTAAAAATAAAAAATTAAGAAATGTTTATACAGTTGATAAAGATCTATTATA 834
 QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
 Db 835 AGATTACCTAGTATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 11

AAAS3726
 ID AAAS3726 standard; DNA; 876 BP.

XX AC AAAS3726;

XX DT 15-SEP-2003 (revised)

XX DT 22-DEC-2000 (first entry)

XX DE Campylobacter jejuni O:19 serotype CstII sialyltransferase.

XX DE Biosynthetic locus; biosynthesis; lipid A biosynthesis;

XX KW acetyltransferase; glycosyltransferase; Beta-1,4-GalNAc transferase;

XX KW Beta-1,3-galactosyltransferase; alpha-2,3-sialyltransferase;

XX KW sialic acid synthase; CMP-sialic acid synthetase; mimic; antibody;

XX KW immunity; immunogen; ganglioside; ds.

XX OS Campylobacter jejuni; O:19 serotype.

XX FH Key Location/Qualifiers

XX FT CDS 1..876

XX FT /*cag= a

XX FT /product= "CstII sialyltransferase"

XX XX WO200046379-A1.

XX XX 10-AUG-2000.

XX XX 01-FEB-2000; 2000WO-CA000086.

XX XX 01-FEB-1999; 99US-0118213P.

XX XX 31-JAN-2000; 2000US-00495406.

XX XX (CANA) NAT RES COUNCIL CANADA.

XX XX Gilbert M, Wakarchuk WW;

XX XX WPI; 2000-524418/47.

XX XX P-PSDB; AAY97212.

XX XX Novel glycosyltransferase polypeptides and polynucleotides useful for

XX XX biosynthesis of ganglioside and ganglioside mimics, as diagnostic

XX XX reagents and as immunogen for producing antibodies.

XX XX Disclosure; Page 96; 120pp; English.

XX XX A reaction mixture for the synthesis of a sialylated oligosaccharide is

XX XX useful for synthesising sialylated oligosaccharide such as ganglioside,

XX XX lysoganglioside or their mimics. Glycosyltransferases are useful for

CC chemo-enzymatic synthesis of oligosaccharides, including gangliosides and
 CC other oligosaccharides that have biological activity. The enzymes and
 CC nucleic acids that encode them are useful for studies of the pathogenesis
 CC mechanisms of organisms that synthesize ganglioside mimics, such as C.
 CC jejuni and the nucleic acids are used as probes to study expression of
 CC genes involved in ganglioside mimetic synthesis. Antibodies raised
 CC against the glycosyltransferases are also useful for analyzing the
 CC expression patterns of these genes involved in pathogenesis. The nucleic
 CC acids are also useful for designing antisense oligonucleotides for
 CC inhibiting expression of the Campylobacter enzymes that are involved in
 CC the biosynthesis of ganglioside mimics that can mask the pathogens from
 CC the host's immune system. The oligosaccharides are useful as diagnosing
 CC reagents or as therapeutics and as immunogens for producing antibodies.
 CC Bacterial glycosyltransferase can be used to catalyse the formation of
 CC oligosaccharides that are identical to the corresponding mammalian
 CC structures and are easier and less expensive to produce in large
 CC quantity, compared to the mammalian glycosyltransferase. The bacterial
 CC origin of the enzymes facilitates expression of large quantities of the
 CC enzymes using relatively inexpensive prokaryotic expression systems.
 CC (updated on 15-SEP-2003 to standardise OS field)

XX SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.8e-71 Length: 876
 Score: 804.00 Matches: 158
 Percent Similarity: 69.28% Conservative: 45
 Best Local Similarity: 53.92% Mismatches: 72
 Query Match: 51.18% Indels: 18
 DB: 3 Gaps: 6

US-10-735-419-10 (1-294) x AAAS3726 (1-876)

QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
 Db 13 ATATTGCTGGAATGGACCAAGTTTAAAGAAATGATTATTCAAGGCTACCAATGAT 72
 QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db 73 TTTGATGATATTAGATGTAATCAATTTTATTGTAAGATAAAATACTATCTCTGTAATAAAA 132
 QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
 Db 133 TGCAAAGCAGTGTTTTACACCCCTAATTTCTTTGAGCAATACTACACTTTTAAACAAT 192
 QY 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 Db 193 TTAATCCAAATCAAGATATGAGCCGAACTAAATATGTTCTTAATTAACCAAGCT 252
 QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db 253 CATCTAGAAATGAAAAATTTTGTAAAAACTTTTACGATTATTTCCTCATGCTCATTTG 312
 QY 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
 Db 313 GGATATGATTTTTTTAAACAACACTTAAAGAAATTAATGCTTATTTTAAATTTTCCGAAAT 372
 QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 Db 373 TATTTCATCAAGAATTAACCTCAGGGGTCTATATGTGTGTCAGTAGCCATAGCCCTAGGA 432
 QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
 Db 433 TACAAAGAAATTTATCTTTCGGGAATTTGATTTTTTATCAAAATGGGTCACT---TATGCT 489
 QY 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 Db 490 TTTGATACCAACAAAGAAATCTTTTAAACCTAGCCCCCTGATTTTAAAAATGATCGCTCG 549
 QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 Db 550 CACTATATCGCATAGTAAAAATACAGATATAAAGCTTTTGAATTTCTAGAAAAAACT 609

QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAACAATATATGCTTATGCTCCTAATAGTCTTTTAGCAAAATTTTATAGAACA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 CGCCCAATTAATTCNAATTTTATCATACAGAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACCTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
Db 775 AATTTTAAAAAATAAATAAATAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTATATAAGCAATTATTTCAAAGGAAAA 873
RESULT 12
ABT13669
ID ABT13669 standard; DNA; 876 BP.
XX
AC ABT13669;
XX
07-FEB-2003 (first entry)
XX
DE C. jejuni bifunctional sialtransferase cstII coding sequence #4.
XX
KW Gene; ds; enzyme; gene therapy; acyltransferase; glycosyltransferase;
KW GalNAc transferase; N-Acetylgalactosamine transferase;
KW galactosyltransferase; sialyltransferase; sialic acid synthase;
KW cytidine 5'-monophosphate sialic acid synthetase;
KW CMP sialic acid synthetase; acetyltransferase; ganglioside synthesis;
KW ganglioside mimetics; inflammation; tumour metastasis.
XX
OS Campylobacter jejuni.
XX
SN WO200274942-A2.
PN
PD 26-SEP-2002.
XX
PF 22-FEB-2002; 2002WO-CA000229.
PR
XX 21-MAR-2001; 2001US-00816028.
XX
PA (CANA) NAT RES COUNCIL CANADA.
XX
PI Gilbert M, Wakarchuk WW;
XX
DR WPI; 2003-040554/03.
DR P-PSDB; ABJ18482.
XX
PT New glycosyltransferases from Campylobacter, useful for synthesizing
PT gangliosides and ganglioside mimetics, and in studying the pathogenesis
PT mechanisms of organisms that synthesize ganglioside mimetics.
XX
PS Disclosure; Page 98; 107pp; English.
XX
CC The invention comprises the amino acid and coding sequences of
CC Campylobacter jejuni proteins. The C. jejuni proteins of the invention
CC may be either an: acyltransferase; glycosyltransferase; GalNAc (N-
CC Acetylgalactosamine) transferase; galactosyltransferase;
CC sialyltransferase; sialic acid synthase; cytidine 5'-monophosphate (CMP)
CC sialic acid synthetase; acetyltransferase. The C. jejuni DNA and protein
CC sequences of the invention are useful for ganglioside synthesis, studying
CC ganglioside mimetics, and for designing oligonucleotides to inhibit
CC expression of Campylobacter enzymes involved in the biosynthesis of
CC ganglioside mimetics that can mask the pathogen's from the host's immune
CC system. The C. jejuni oligosaccharides of the invention may be used as
CC diagnostic reagents (e.g. to locate areas of inflammation or tumour
CC metastasis). The present DNA sequence represents a Campylobacter jejuni

CC gene of the invention
XX
SQ Sequence 876 BP; 353 A; 117 C; 109 G; 297 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.8e-71 Length: 876
Score: 804.00 Matches: 158
Percent Similarity: 69.28% Conservative: 45
Best Local Similarity: 53.92% Mismatches: 72
Query Match: 51.18% Indels: 18
DB: 8 Gaps: 6
US-10-735-419-10 (1-294) x ABT13669 (1-876)
QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATATTGCTGGAAATGGACCAAGTTTAAAGAAATGATTATTCAGGGTACCAATGAT 72
QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGTATTAGATGTAATCAATTTTATTGAGATAAATACTATCTTGTGTAATAAAA 132
QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCAAAGCAGTGTGTTTACACCCCTAATTTCTTTGAGCAATACTACTTTAAAAACAT 192
QY 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAAATCCAAATCAAGATATGAGCCGAACTAATATGTTGTTCTTAATTACACCAAGCT 252
QY 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAATGAAATTTTGTAAACAACTTTTACGATTATTTTCTCATGCTCATTTG 312
QY 108 GlyHisGluIleIleLysLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGATATGATTTTTTAAACAACTTAAAGAAATTAATGCTTATTTTAAATTTTACGAAAT 372
QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCATCAAGAAATTACCTCAGGGTCTATATGTGTGCGATAGCCATAGCCCTAGGA 432
QY 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTTCCGGCAATGATTTTATCAAAATGGGTCACT---TATGCT 489
QY 168 PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTGATACCAACCAAGAAAAATCTTTAAACACTAGCCCTCGATTTTAAAAAATGATCGCTCG 549
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGGACATAGTAAATAATACAGATATAAAGCTTTAGAAATTTTACAAAAAATCT 609
QY 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAACAATATATGCTTATGCTCCTAATAGTCTTTTAGCAAAATTTTATAGAACA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 CGCCCAATTTTAAATTCNAATTTTATCATACAGAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACCTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuIle 275
Db 775 AATTTTAAAAAATAAATAAATAAAGAAATGTTTATTACAAAGTTGATAAAGATCTATTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTATATAAGCAATTATTTCAAAGGAAAA 873

RESULT 13

AAZ25693
ID AAZ25693 standard; DNA; 1293 BP.XX AC AAZ25693;
XX DT 05-JAN-2000 (first entry)

XX DE Campylobacter jejuni alpha-2,3-sialyltransferase gene.

XX KW Campylobacter jejuni; alpha-2,3-sialyltransferase; cst-I; acceptor;
XX KW lipopolysaccharide; galactose residue; sialic acid molecule; ss.

XX OS Campylobacter jejuni.

XX Key Location/Qualifiers
XX FH 1..1293
XX FT CDS /*tag= a
XX PT

XX PN WO9949051-A1.

XX PD 30-SEP-1999.

XX PF 22-MAR-1999; 99WO-CA000238.

XX PR 20-MAR-1998; 98US-0078891P.

XX PR 18-MAR-1999; 99US-00272960.

XX PA (CANADA) NAT RES COUNCIL CANADA.

XX PI Gilbert M, Wakarchuk WW;

XX DR WPI; 1999-601216/51.

XX DR P-PSDB; AAY45221.

XX PT New sialyltransferases useful for adding sialyl residues to acceptor molecules.
XX PS Claim 7; Fig 2; 47pp; English.
XX CC The present sequence represents the Campylobacter jejuni cst-I gene which encodes alpha-2,3-sialyltransferase. The alpha-2,3-sialyltransferase protein is useful for producing desired carbohydrate structures by contacting the acceptor molecule (which has a terminal galactose residue) with an activated sialic acid molecule. The terminal galactose residue is linked to a second residue (Glc or a GlcNAc or GalNAc) in the acceptor molecule through a beta-1,3 or beta-1,4 linkage, respectively. The activated sialic acid is CMP-Neu5Ac. The polynucleotides and polypeptides facilitate the improved production of desired structures and nucleic acids encoding sialyltransferases

XX SQ Sequence 1293 BP; 515 A; 151 C; 176 G; 451 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.88e-59 Length: 1293
Score: 680.50 Matches: 145
Percent Similarity: 66.44% Conservative: 49

Best Local Similarity: 49.66% Mismatches: 75

Query Match: 43.32% Indels: 23

DB: 2 Gaps: 8

US-10-735-419-10 (1-294) x AAZ25693 (1-1293)

QY 1 MetSerMetAsnIle---AsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIle 19

DB 34 GTTAGTAAATAATGCAAAATATATATCATGAGGAAATGGACCTAGCTTAAATAATATT 93

QY 20 AsPTyTlysArgLeuProLysGlnPheAspValPheArgCysAsnGlnPheTyrPheGlu 39

DB 94 AATTATAAAGACGTGCTAGAGATATGATGTTTTTAGGTGTAACACGTTTTTTTGAA 153

QY 40 AspArgTyrPheValGlyLysAspValLysTyrValPheAsnProPheValPhePhe 59

Db 154 GATAAGTATTATTAGGAAAAAAGATTAAAGCAGTATTTTAACTCGTGCTTTTAA 213

QY 60 GluGlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIle 79

Db 214 CAACAGTATCACACTGCAAAACAACATTACTAAAAAATGAGTATGAAATAAAAAATATT 273

QY 80 ValCysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGlu 99

Db 274 TTTTGCTCTACATTTAAATTTTACCTTTTGAAGCAATGATTTTTTACATCAATTTAT 333

QY 100 LeuTyrPheSerAspAlaPheLeuGlyHisGluIleLeuLysLysLeuLysAspPhePhe 119

Db 334 AATTTTTCCTCCGATGCAAAACTTGGCTATGAGTTATTGAAAACCTTAAAGAAATTTAT 393

QY 120 AlaTyrIleLysTyrAsnGluIleTyrAsnA-gGlnArgIleThrSerGlyValTyrMet 139

Db 394 GCTTATATAAATAACAATGAAATTTATTCAATAAAGAATTACTTCGGCGCTATATG 453

QY 140 CysAlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyr 159

Db 454 TGTGCAATTTGCTATTGTCATTAGGATATAAAACCATCTATTATTATGTGCGCATTTTAT 513

QY 160 GlnAspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysAsnLeuLysCys 179

Db 514 GAAGGA---GATGTTATTTATCTTTTGAAGCTAGTACAAATATATAAAACAACTTTT 570

QY 180 ThrGlyPheLysAsnGlnLysPheLysPheIleAsn---HisSerMetAlaCysAspLeu 198

Db 571 CTGGAATAAAGAT-----TTCAACCTTCAAAATTTGTCATTCTAAGGAATACGATATA 624

QY 199 GlnAlaLeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAsp 218

Db 625 GAAGCATTTAAATTTGTTAAATCAATATATACAAAGTTAATATCTACGCATTTGTGTGATG 684

QY 219 -----GluTyrPheLysLeuAlaProAspIleGlySerAspPheValLeuSer 234

Db 685 TCTATTTTGGCAATCAITTTCTTATCAATTAATTAATTAATTAATTAATTTTAA 744

QY 235 LysLysProLysLysTyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArg 254

Db 745 AATAAGCAATAAATTTCTATAATGATATTTTATTGACTGATAATACTCTCGCGTAAGT 804

QY 255 TyrTyrGlyLysSerArgLeuLysGluAsn----- 265

Db 805 TTTTAT-----AAAATCAACTTAAAGCTGATAATAAAATTAATGCTTAATTTTATAAT 858

QY 266 ---LeuHisTyrLys-----LeuIleLysAspLeu 274

Db 859 ATTCTTCATCTTAAGATAATTTAAATTAATTTTAA 894

RESULT 14

ACCT71694

ID ACC71694 standard; DNA; 909 BP.

XX AC ACC71694;

XX DT 23-JUL-2003 (first entry)

XX DE alpha-2,3/alpha-2,8-sialyltransferase coding sequence #3.

XX KW alpha-2,3/alpha-2,8-sialyltransferase; enzyme; sialic acid; gene; ds.

XX OS Pasteurella multocida.

XX Key Location/Qualifiers

XX FH 1..909 a

XX FT /*tag= a

XX FT /product= "alpha-2,3/alpha-2,8-sialyltransferase #3"

XX FT /note= "No stop codon given"

XX PN WO2003027297-A1.

XX


```
DB: 13 Gaps: 6
US-10-735-419-10 (1-294) x ADT05647 (1-191996)
Qy 6 AsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuPro 25
Db 43856 TCTGTCATTATGTCAGGTAAATGGAACAAGTTTAAATCAATTGACTATAGTTTATACCT 43797
Qy 26 LysGlnPheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGly 45
Db 43796 AAAGATTATGATGTTTCCGTTGCAATCAATTTATTTTGGAGTCAATTTTCTTGGT 43737
Qy 46 LysAspValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSer 65
Db 43736 AAGAAATATAAAAGGTATTTTAAATGTTCTGTAATTTTGAACAATACTATACGTTT 43677
Qy 66 LysLysLeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThr---Ile 84
Db 43676 ATGCAATTAAATAAATAATGAATATGAATATGCTGATATTTATTCATCTTTCTA 43617
Qy 85 AsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAsp 104
Db 43616 AATTAGGGATTTCAGAA--TTAAGAAATCCAGCGTTTGAAGAAATTACTACCACAA 43560
Qy 105 AlaPheLeuGlyHisGluIleLysLysLeuLysAspPhePheAlaTyrIleLysTyr 124
Db 43559 ATCGATCTTGTCATAGCTATTTAAATAAACTACGAGCTTTTGATGCTCAATTTACAAT 43500
Qy 125 AsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaVal 144
Db 43499 CACGAATATATGAGATAAGAGGATTACATCAGCGCTCTATATGTCGAGTGGCACT 43440
Qy 145 AlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsn 164
Db 43439 GCAATGGGTTATAAAGATCTTTATTTCACAGCATTCATTTTATCAAGAAAGGGAAT 43380
Qy 165 LeuTyrAlaPheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsn 184
Db 43379 CTTTACCGATTTTCATCAACAGAAATATTTAAATTTATTAATTTACCTCTTTTTCACAA 43320
Qy 185 GlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeu 204
Db 43319 AATAAAGTCAAGCGATATCCATTCTATGGAATATGATTTAATGCACTTTATTTTAA 43260
Qy 205 MetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAsp-----GluTyr 220
Db 43259 CAAAACATTATGGAGTAATATTTATTGCAATTCGCCAGAAAGTCCTCTATGTAATTAT 43200
Qy 221 PheLysLeuAlaPro---AspIleGlySerAspPheValLeuSerLysLysProLysLys 239
Db 43199 TTCCCTTTATCACCACCTGAATAACCCCAATTACTTTTATTCGAGAAAG--AAAAAT 43143
Qy 240 TyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyr---TyrGlyLys 258
Db 43142 TACACACAGATATTTTAAATTCGCCGAGTTTGATATATAAATAATTTGGTATATATCC 43083
Qy 259 LysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIleArgLeuPro 278
Db 43082 AAACCAAGATTTACCAAAATCTGATTTTCGGTTGATCTGGGATATATTACGTTTACCT 43023
Qy 279 SerAspIleLysHisTyrLeuLysGluLys 288
Db 43022 AATGATATAAACACGCTTTGAAAGCAAG 42993
```

Search completed: June 14, 2005, 18:49:07
Job time : 692 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 18:24:29 ; Search time 204 Seconds
(without alignments)
2358.164 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSNNALVCGNPSLKNID.....IRLPDIKHVLYKEKYANKNR 294

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10735419/runat_13062005_165725_12682/app_query.fasta_1.455
-DB=Issued Patents NA -OPMT=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosun62 -TRANS=human40.cdi
-LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10735419 @CGN 1 1 177 @runat_13062005_165725_12682 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	823	52.4	876	4	US-09-495-406-2
2	823	52.4	876	4	US-09-816-028A-2
3	823	52.4	876	4	US-10-303-162-2
4	823	52.4	876	4	US-10-303-134-2
5	823	52.4	11474	4	US-09-495-406-1
6	823	52.4	11474	4	US-09-816-028A-1
7	823	52.4	11474	4	US-10-303-162-1
8	823	52.4	11474	4	US-10-303-134-1
9	817	52.0	876	4	US-09-495-406-6
10	817	52.0	876	4	US-09-816-028A-6
11	817	52.0	876	4	US-10-303-162-6
12	817	52.0	876	4	US-10-303-134-6

13	813	51.8	876	4	US-09-495-406-4	Sequence 4, Appli
14	813	51.8	876	4	US-09-816-028A-4	Sequence 4, Appli
15	813	51.8	876	4	US-10-303-162-4	Sequence 4, Appli
16	813	51.8	876	4	US-10-303-134-4	Sequence 11, Appl
17	804	51.2	873	4	US-09-816-028A-11	Sequence 11, Appl
18	804	51.2	873	4	US-09-816-028A-13	Sequence 11, Appl
19	804	51.2	873	4	US-10-303-162-11	Sequence 13, Appl
20	804	51.2	873	4	US-10-303-162-13	Sequence 13, Appl
21	804	51.2	873	4	US-10-303-134-11	Sequence 11, Appl
22	804	51.2	873	4	US-10-303-134-13	Sequence 8, Appli
23	804	51.2	876	4	US-09-495-406-8	Sequence 8, Appli
24	804	51.2	876	4	US-09-816-028A-8	Sequence 8, Appli
25	804	51.2	876	4	US-10-303-162-8	Sequence 8, Appli
26	804	51.2	876	4	US-10-303-134-8	Sequence 1, Appli
27	680.5	43.3	1293	4	US-09-272-960-1	Sequence 1, Appli
28	680.5	43.3	1293	4	US-10-058-636-1	Sequence 1, Appli
C 29	662.5	42.2	1830121	4	US-09-557-884-1	Sequence 1, Appli
C 30	662.5	42.2	1830121	4	US-09-643-990A-1	Sequence 1, Appli
C 31	114	7.3	640681	4	US-09-790-988-1	Sequence 20, Appl
32	113.5	7.2	3270	4	US-09-710-279-4353	Sequence 4353, Ap
33	113	7.2	21338	3	US-08-561-527-20	Sequence 1, Appli
C 34	112.5	7.2	640681	4	US-09-790-988-1	Sequence 1, Appli
35	112	7.1	3510	4	US-09-583-110-1748	Sequence 1748, Ap
36	112	7.1	3510	4	US-09-107-433-2244	Sequence 2244, Ap
37	110	7.0	1664976	4	US-08-916-421B-1	Sequence 1, Appli
38	110	7.0	1664976	4	US-09-692-570-1	Sequence 1, Appli
39	109	6.9	4344	4	US-09-601-198-165	Sequence 165, App
C 40	108.5	6.9	1664976	4	US-08-916-421B-1	Sequence 1, Appli
C 41	108.5	6.9	1664976	4	US-09-692-570-1	Sequence 1, Appli
C 42	108	6.9	1712	3	US-09-144-7768-7	Sequence 7, Appli
43	108	6.9	1712	4	US-08-882-431B-7	Sequence 1, Appli
C 44	106	6.7	580073	4	US-08-545-528B-1	Sequence 1, Appli
45	105.5	6.7	1101	4	US-09-861-451A-45	Sequence 45, Appl

ALIGNMENTS

RESULT 1

US-09-495-406-2
; Sequence 2, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)

OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
OTHER INFORMATION: strain OH4384 (ORF 7a of LOS biosynthesis locus)
US-09-495-406-2

Alignment Scores:
Pred. No.: 1.1e-92
Score: 823.00
Percent Similarity: 69.62%
Best Local Similarity: 54.95%
Query Match: 52.39%
DB: 4
Length: 876
Matches: 161
Conservative: 43
Mismatch: 71
Indels: 18
Gaps: 6


```

Db      610 TACAAATAAACATATATGCTTATGCTTACAGCTCTTTTAGCAAAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLysLeuLysAspLeuIle 275
Db      775 AATTTTAAAAATAAATAATTAAGAAATATTTATTTACAGTTTGATTAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACTAGTATATAAGCAATTATTTCAAGGAAAA 873

RESULT 3
US-10-303-162-2
; Sequence 2, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-303-162-2

Alignment Scores:
Pred. No.: 1,1e-92 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 4 Gaps: 6

US-10-735-419-10 (1-294) x US-10-303-162-2 (1-876)
Qy      8 LeuValCysGlyAangLysProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db      13 ATATTGCTGGAATGACCAAGTTTAAAGAAATGATTATTTCAAGACTACCAATGAT 72
Qy      28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGTTTATGATGTAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAA 132
Qy      48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67

```

```

Db      133 TGCAGGCGAGTATTTTACAACTCTATCTCTTTTGTGAACAATACACTACACTTTAAAAACAT 192
Qy      68 LeuIleGlnAsnGluGlyTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db      193 TTAATCCAAATCAAGAAATATGAGACGGAATAATTATGTTGTTCTTAATTACCAACCAAGCT 252
Qy      88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAATAAGAAATTTGTAAACACTTTTACGATTATTTTCTCGTAGCTCATTTG 312
Qy      108 GlyHisGluIleIleLysLysLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db      313 GGATATGATTTTTCACCAACTTAAGATTTTAAGCTTATTTTAATTTTACGAATTT 372
Qy      128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAATACTACCTCAGGGGTTTATATGTGTGCAGTAGCATAGCCCTAGGA 432
Qy      148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCGGGAATGATTTTATCAAAATGGGTCTCT---TATGCT 489
Qy      168 PheAspAsnAsnLysLysAsnLeuLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490 TTTGATACTAAACAAAAAATCTTTTAAATTCGCTCTTAATTTTAAATATGATAATCA 549
Qy      188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGACATAGTAAAAATACAGATATATAAAGCTTTAGAAATTTCTAGAAAAA 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAATAAAACATATATGCTTATGCTTACACAGCTTTTATAGCAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTTCAAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLysLeuLysAspLeuIle 275
Db      775 AATTTTAAAAATAAATAATTAAGAAATATTTTATTTACAGTTTGATTAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACTAGTATATAAGCAATTATTTCAAGGAAAA 873

RESULT 4
US-10-303-134-2
; Sequence 2, Application US/10303134
; Patent No. 6825019
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; TITLE OF INVENTION: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876

```



```

Db      6480  TACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 6536
Qy      168   PheAspAenAenLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      6537  TTTGATACTAAACAACAAATAATCTTTTAAATTTGGCTCCTCAATTTTAAATAATGATAATTCA 6596
Qy      188   LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      6597  CACTATATCGGACACTAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 6656
Qy      208   TyrAspValAenIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      6657  TACAAATAAACAATATATGCTTATGCTTACAGCTCTTTAGCAAAATTTTATAGAACTA 6716
Qy      224   AlaProAspIleGlySerAspPheValLysSerLysLysProLysLysTyrIleAsnAsp 243
Db      6717  GCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAA---AATAACTACACTAAAGAT 6773
Qy      244   IleLeuIleProAspLysTyrAlaGlnGluArgTyrGly----- 257
Db      6774  ATACTCATACCTCTAGTGGCT-----TATGCAAAATTTTCAAAAAATATT 6821
Qy      258   -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeu 275
Db      6822  AATTTTAAAAAATAAATAATTAAGAAATATTTATTTACAGTTGATAAAGATCTATTA 6881
Qy      276   ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      6882  AGATTACCTAGTATATAAGCATTTATTTCAAGGAAAA 6920

```

RESULT 6

```

US-09-816-028A-1
; Sequence 1, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 11474
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
; OTHER INFORMATION: including LOS biosynthesis locus
US-09-816-028A-1

```

Alignment Scores:

```

Pred. No.: 5.37e-91 Length: 11474
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 4 Gaps: 6

```

US-10-735-419-10 (1-294) x US-09-816-028A-1 (1-11474)

```

Qy      8   LeuValCysGlyAenGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db      6060  ATATTCTGGAATGGACCAAGTTTAAAGAAATTTATTTACAGACTACCAATGAT 6119

```

```

Qy      28   PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      6120  TTTGATGTATTTAGATGTAACTCAATTTTATTTTGAAGATAAAATACTCTTGTGTAATAAAA 6179
Qy      48   ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db      6180  TGCNAGGCAGTATTTTACAATCTTATCTTTTGTGTTCTAATTTACAACTATACATCTTTAAACAT 6239
Qy      68   LeuIleGlnAsnGluTyrAenIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db      6240  TTAATCCAAATCAAGAATATGAGACCGAACTAATATATGTGTCTAATTTACAACTCAAGCT 6299
Qy      88   TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      6300  CATCTAGAAAATGAAAATTTTGTAAACCTTTTTACGATTATTTTCTGTGATGCTCATTTG 6359
Qy      108  GlyHisGluIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db      6360  GGATATGATTTTTCACCAACTTAAGATTTTAATGCTTATTTTAAATTTTCAGAAATTT 6419
Qy      128  TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      6420  TATTTCAATCAAGAATATACCTCAGGGGTCTATATGTGTGTCAGTAGCCATAGCCCTAGGA 6479
Qy      148  TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnLeuTyrAla 167
Db      6480  TACAAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 6536
Qy      168  PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      6537  TTTGATACTAAACAACAAATAATCTTTTAAATTTGGCTCCTTAATTTTAAATGATAATTCA 6596
Qy      188  LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      6597  CACTATATCGGACACTAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 6656
Qy      208  TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      6657  TACAAATAAACAATATATGCTTATGCTTATCTCAACAGCTCTTTTACGAAATTTTATAGAACTA 6716
Qy      224  AlaProAspIleGlySerAspPheValLysSerLysLysProLysLysTyrIleAsnAsp 243
Db      6717  GCGCCAAATTTAAATTTCAATTTTATCATCAAGAAAA---AATAACTACACTAAAGAT 6773
Qy      244  IleLeuIleProAspLysTyrAlaGlnGluArgTyrGly----- 257
Db      6774  ATACTCATACCTCTAGTGGCT-----TATGCAAAATTTTCAAAAAATATT 6821
Qy      258  -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeu 275
Db      6822  AATTTTAAAAAATAAATAATTAAGAAATATTTATTTACAGTTGATAAAGATCTATTA 6881
Qy      276  ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      6882  AGATTACCTAGTATATAAGCATTTATTTCAAGGAAAA 6920

```

RESULT 7

```

US-10-303-162-1
; Sequence 1, Application US/10303162
; Patent No. 6723545
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,162
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01

```

Db 6774 ATACTCATACCTTCTAGTAGGCT-----TATGGAAAAATTTTCAAAAAATATT 6821
 Qy 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuLys 275
 Db 6822 AATTTTAAAAAATAAAAAATAAGAAAAATATTATTACAAGTTGATAAAGATCTATTA 6881
 Qy 276 ArgLeuProSerAspLysHisTyrLeuLysGluLys 288
 Db 6882 AGATTACTAGTATATAAGCATTAATTTTCAAGGAAAA 6920
 RESULT 8
 US-10-303-134-1
 ; Sequence 1, Application US/10303134
 ; Patent No. 6825019
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: Wakarchuk, Warren W.
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/303,134
 ; CURRENT FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US/09/816,028
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 11474
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; OTHER INFORMATION: 11.5 kb PCR product from C. jejuni OH4384
 ; OTHER INFORMATION: including LOS biosynthesis locus
 US-10-303-134-1
 Alignment Scores:
 Pred. No.: 5,37e-91 Length: 11474
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Best Local Similarity: 54.95% Mismatches: 71
 Query Match: 52.39% Indels: 18
 DB: 4 Gaps: 6
 US-10-735-419-10 (1-294) x US-10-303-134-1 (1-11474)
 Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
 Db 6060 ATTATTGCTGGAATGGACCAGTTTAAAGAAATTTGATTATTCAGAGCTACCACAATGAT 6119
 Qy 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db 6120 TTTTGATGTTATTAGATGTAATCAATTTATTTTGAAGATAAATACTATCTTGGTAAAAA 6179
 Qy 48 ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
 Db 6180 TGCAAGGCGAGTATTTTACAACTCCTATTTCTTTTTTGAACAATACTACACTTTAAAAACAT 6239
 Qy 68 LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 Db 6240 TTAATCCAAATCAAGAAATAGACCGAACTTAATATTATGTGTTCTAATTACAACCAAGCT 6299
 Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db 6300 CATCTAGAAAAAGAAAATTTTGTAAAAAATTTTTCAGATTATTTTCTGATGTCATTG 6359
 Qy 108 GlyHisGluIleLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluLys 127
 Db 6360 GGATATGATTTTTTCAACAACACTTAAAGATTTTAAATGCTTTATTTTAAATTTTCAGAAAT 6419

QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
|||
Db 6420 TATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGTGTCAGTAGCCATAGCCCTAGGA 6479
|||
QY 148 TyrlsSerIleTyrlsSerGlyIleAspPheTyrlsGlnAspThrAsnLeuTyrAla 167
|||
Db 6480 TACAAAGAAATTTATCTCTTCGGGAATTTGATTTTATCAAAATGGGTATCT---TATGCT 6536
|||
QY 168 PheAspAsnLeuLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
|||
Db 6537 TTTTGATACTAAACAAACAAATCTTTTAAATTTGGCTCTCTTAATTTTAAATGATAATTCA 6596
|||
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
|||
Db 6597 CACTATATCGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 6656
|||
QY 208 TyrlsPheValAsnIleTyrlsSerLeu---AsnSer---AspGluTyrPheLysLeu 223
|||
Db 6657 TACAAATAAAACATATATGCTTATGTCTTAACAGTCTTTTACAAATTTTATAGAACTA 6716
|||
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrlsTyrlsLeuAsnAsp 243
|||
Db 6717 GCGCCAAATTTAAATTCATATTTATCATACAGAAAA---AATAACTACACTAAAGAT 6773
|||
QY 244 IleLeuIleProAspLysTyrlsAlaGlnGluArgTyrTyrlsGly----- 257
|||
Db 6774 ATACTCATACCTTCTAGTGAGGCT-----TATGGAATAATTTTCAAAAAATATT 6821
|||
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrlsLysLeuIleLysAspLeuIle 275
|||
Db 6822 AATTTTAAAAATAAATAAATAAAGAAATATTATTACAAAGTTGATAAAGATCTATTATTA 6881
|||
QY 276 ArgLeuProSerAspIleLysHisTyrlsLysLeuLysGluLys 288
|||
Db 6882 AGATTACCTAGTATAGACATATTATTCAGAGAAAA 6920
|||

RESULT 9

US-09-495-406-6
; Sequence 6, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; CURRENT FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(876)
; OTHER INFORMATION: Campylobacter alpha-2,3/alpha 2,8-sialyltransferase II
; OTHER INFORMATION: (cstII) from C. jejuni serotype O:41
US-09-495-406-6

Alignment Scores:

Pred. No.:	6.12e-92	Length:	876
Score:	817.00	Matches:	159
Percent Similarity:	69.62%	Conservative:	45
Best Local Similarity:	54.27%	Mismatches:	71
Query Match:	52.01%	Indels:	18
DB:	4	Gaps:	6

US-10-735-419-10 (1-294) x US-09-495-406-6 (1-876)

QY 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
|||
Db 13 ATTATTGCTGGAATTCGACCAAGTTTAAAGAAATTTGATTTTCAAGACTTACCAAAATGAT 72
|||
QY 28 PheAspValPheArgCysAsnGlnPheTyrlsPheGluAspArgTyrPheValGlyLysAsp 47
|||
Db 73 TTTGATGTTATTTAGATGCAATCAATTTTATTTGAAGATAAATACTATCTTTGGTAAAAA 132
|||
QY 48 ValLysTyrlsValPhePheAsnProPheValPhePheGluGlnTyrTyrlsSerLysLys 67
|||
Db 133 TGCAGAGCAGTATTTTACATCTCTTTTGTGAACAATACTTACACTTTTAAAAACAT 192
|||
QY 68 LeuIleGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
|||
Db 193 TTAATCCAAAATCAAGATATGAGACCGAACTAATCATGTGTTCTTAATTTTAAACCAAGCT 252
|||
QY 88 TyrlsAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
|||
Db 253 CATCTAGAAAATCAAAATTTTGTAAAACTTTTACGATTATTTTCTGATGCTCATTTG 312
|||
QY 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrlsTyrlsAsnGluIle 127
|||
Db 313 GGATATGATTTTTCACCAACTTAAGAAATTCATGCTTATTTTAAATTTTACAGAAAT 372
|||
QY 128 TyrlsAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
|||
Db 373 TATTTCAATCAAGAATTAACCTCAGGGGTCTATATGTGCACAGTAGCCATAGCCCTAGGA 432
|||
QY 148 TyrlsSerIleTyrlsSerGlyIleAspPheTyrlsGlnAspThrAsnLeuTyrAla 167
|||
Db 433 TACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGATCATCT---TATGCT 489
|||
QY 168 PheAspAsnAsnLysLysAsnLeuLysCysThrGlyPheLysAsnGlnLysPhe 187
|||
Db 490 TTTGATACCAACAAACAAATCTTTTAAATTTGGCTCTCTTAATTTTAAATGATAATTCA 549
|||
QY 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
|||
Db 550 CACTATATCGCATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
|||
QY 208 TyrlsPheValAsnIleTyrlsSerLeu---AsnSer---AspGluTyrPheLysLeu 223
|||
Db 610 TACGAAATAAAGCTATATTTGTTATGCTCTTAACAGCTTTTATAGCAATTTTATAGAACTA 669
|||
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrlsTyrlsAsnAsp 243
|||
Db 670 GCGCCAAATTTAAATTCATATTTATCATACAGAAAA---AATAACTATACTAAAGAT 726
|||
QY 244 IleLeuIleProAspLysTyrlsAlaGlnGluArgTyrTyrlsGly----- 257
|||
Db 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
|||
QY 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrlsLysLeuIleLysAspLeuIle 275
|||
Db 775 AATTTTAAAAATAAATAAATAAAGAAATATTATTACAAAGTTGATAAAGATCTATTATTA 834
|||
QY 276 ArgLeuProSerAspIleLysHisTyrlsLysGluLys 288
|||
Db 835 AGATTACCTAGTATAGACATATTATTCAGAGAAAA 873
|||

RESULT 10

US-09-816-028A-6
; Sequence 6, Application US/09816028A
; Patent No. 669705
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US


```

RESULT 13
US-09-495-406-4
; Sequence 4, Application US/09495406
; Patent No. 6503744
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-0001110US
; CURRENT APPLICATION NUMBER: US/09/495,406
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of LOS biosynthesis locus)
US-09-495-406-4

Alignment Scores:
Pred. No.: 1,92e-91 Length: 876
Score: 813.00 Matches: 159
Percent Similarity: 69.97% Conservative: 46
Best Local Similarity: 54.27% Mismatches: 70
Query Match: 51.75% Indels: 18
DB: 4 Gaps: 6

US-10-735-419-10 (1-294) x US-09-495-406-4 (1-876)

QY 8 LeuValCysGlyAenGlyProSerLeuLysAenLeuLysAenLeuLysAenLeuLysGln 27
Db 13 ATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGCTACCAATGAT 72

QY 28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGATTTAGATCAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAA 132

QY 48 ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 133 TTCAAAGCAGTATTTTACATCTCGTCTTTTGTGAAACAACTACTACACTTTAAACAT 192

QY 68 LeuLeuGlnAenGluTyrAenLeuGluAenLeuValCysSerThrLeuAenLeuGlu 87
Db 193 TTAATCCAAATCAAGAAATGAGACCAAGTAAATATGTTGTTCTAATTAACACCAAGCT 252

QY 88 TyrLeuAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTCTGTGCTCATTTG 312

QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPheAlaTyrLeuLysTyrAenGluLeu 127
Db 313 GGATATGATTTTTTAAACAACTAAAGCAATTTTAAATGCTTTATTTTAAATTTTCAAGAAAT 372

QY 128 TyrAsnArgGlnArgLeuThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATCTCAATCAAGAAATTTACCTCAGGAGTCTATGTTGTCAGTACTATAGCCCTAGGA 432

QY 148 TyrLysSerLeuTyrLeuSerGlyLeuAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAATGGGTCATCT--TATGCT 489

QY 168 PheAspAenAenLysLysAenLeuLeuAenLysCysThrGlyPheLysAenGlnLysPhe 187

```

```

Db 490 TTTGATACCAACACAGAAAATCTTTTAAAAACTGGCTCCTGATTTTAAAAATGATCGCTCA 549
QY 188 LysPheileAenHisSerMetalCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGCATAGTAAATAAATACAGATATATAAAGCTTTTGAATTTCTAGAAAAAACT 609
QY 208 TyrAspValAenLeuTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAAATAAAACTATATTGCTTATGCTCCTAACAGCTTTTAGCAAAATTTTATAGAACYA 669
QY 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrLeuAenAsp 243
Db 670 GCGCCAAATTTAAATTTTAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
QY 244 IleLeuLeuProAspLysTyrAlaGlnGluAArgTyrTyrGly----- 257
Db 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAAAATTTTCAAAAAATATT 774
QY 258 -----LysLysSerArgLeuLysGluAenLeuHisTyrLysLeuLysAspLeuLeu 275
Db 775 AATTTTAAAAATAAATAAATAAAGAAATATTATTACAGTTGATATAAAGATCTATTATTA 834
QY 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTCAVATAAAGCATTATTTCAAAAGGAAAA 873

RESULT 14
US-09-816-028A-4
; Sequence 4, Application US/09816028A
; Patent No. 6699705
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: Gilbert, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; FILE REFERENCE: 019633-0001110US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR FILING DATE: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: serotype O:10 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-09-816-028A-4

Alignment Scores:
Pred. No.: 1,92e-91 Length: 876
Score: 813.00 Matches: 159
Percent Similarity: 69.97% Conservative: 46
Best Local Similarity: 54.27% Mismatches: 70
Query Match: 51.75% Indels: 18
DB: 4 Gaps: 6

US-10-735-419-10 (1-294) x US-09-816-028A-4 (1-876)

QY 8 LeuValCysGlyAenGlyProSerLeuLysAenLeuLysAenLeuLysGln 27
Db 13 ATTATGCTGGAATGGACCAAGTTTAAAGAAATGATTTTCAAGCTACCAATGAT 72

QY 28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTTGATGATTTAGATCAATCAATTTTATTTTGAAGATAAATACTATCTTGGTAAAAA 132

QY 48 ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 133 TTCAAAGCAGTATTTTACATCTCGTCTTTTGTGAAACAACTACTACACTTTAAACAT 192

QY 68 LeuLeuGlnAenGluTyrAenLeuGluAenLeuValCysSerThrLeuAenLeuGlu 87
Db 193 TTAATCCAAATCAAGAAATGAGACCAAGTAAATATGTTGTTCTAATTAACACCAAGCT 252

QY 88 TyrLeuAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTTTCTGTGCTCATTTG 312

QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPheAlaTyrLeuLysTyrAenGluLeu 127
Db 313 GGATATGATTTTTTAAACAACTAAAGCAATTTTAAATGCTTTATTTTAAATTTTCAAGAAAT 372

QY 128 TyrAsnArgGlnArgLeuThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATCTCAATCAAGAAATTTACCTCAGGAGTCTATGTTGTCAGTACTATAGCCCTAGGA 432

QY 148 TyrLysSerLeuTyrLeuSerGlyLeuAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTTCTGGAATTTGATTTTATCAAAATGGGTCATCT--TATGCT 489

QY 168 PheAspAenAenLysLysAenLeuLeuAenLysCysThrGlyPheLysAenGlnLysPhe 187

```



```
Db      670 GCGCCAAATTAAATTCAAATTTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAAAAATTAAGAAAAATATTATTACAAGTTTGATAAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACCTAGTGATATAAAGCATTTATTTCAAAGGAAAA 873
```

Search completed: June 14, 2005, 21:15:22
Job time : 227 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 20:11:33 ; Search time 671 Seconds
(without alignments)
2716.131 Million cell updates/sec

Title: US-10-735-419-10

Perfect score: 1571

Sequence: 1 MSNNINALVCGNGPSLKNIID.....IRLPDIKHYLKEKYANKNR 294

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 6046767 seqs, 3099530249 residues

Total number of hits satisfying chosen parameters: 12093534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10735419/runat_13062005_165727_12771/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SURF=mrnpb -MINMATCH=0.1
-LOOPCL=0 -LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10735419 @CIGN 1 1 725 @runat_13062005_165727_12771
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US10J_PUBCOMB.seq:*
23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq:*
24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	823	52.4	876	9	US-09-816-028A-2	Sequence 2, Appli
2	823	52.4	876	15	US-10-303-161-2	Sequence 2, Appli
3	823	52.4	876	16	US-10-303-118-2	Sequence 2, Appli
4	823	52.4	876	16	US-10-303-128-2	Sequence 2, Appli
5	823	52.4	876	16	US-10-303-134-2	Sequence 2, Appli
6	823	52.4	876	16	US-10-303-162-2	Sequence 2, Appli
7	823	52.4	876	19	US-10-735-419-2	Sequence 2, Appli
8	823	52.4	876	20	US-10-820-536-2	Sequence 2, Appli
9	823	52.4	876	20	US-10-845-408-2	Sequence 2, Appli
10	823	52.4	876	20	US-10-845-412-2	Sequence 2, Appli
11	823	52.4	876	20	US-10-846-219-2	Sequence 2, Appli
12	823	52.4	876	20	US-10-821-604-2	Sequence 2, Appli
13	823	52.4	876	20	US-10-847-983-2	Sequence 2, Appli
14	823	52.4	876	20	US-10-821-573-2	Sequence 2, Appli
15	823	52.4	876	20	US-10-850-807-2	Sequence 2, Appli
16	823	52.4	876	20	US-10-850-125-2	Sequence 2, Appli
17	823	52.4	876	20	US-10-830-825-2	Sequence 2, Appli
18	823	52.4	876	21	US-10-962-334-2	Sequence 2, Appli
19	823	52.4	876	21	US-10-830-997-2	Sequence 2, Appli
20	823	52.4	876	21	US-10-962-235-2	Sequence 2, Appli
21	823	52.4	11474	9	US-09-816-028A-1	Sequence 1, Appli
22	823	52.4	11474	15	US-10-303-161-1	Sequence 1, Appli
23	823	52.4	11474	16	US-10-303-118-1	Sequence 1, Appli
24	823	52.4	11474	16	US-10-303-128-1	Sequence 1, Appli
25	823	52.4	11474	16	US-10-303-134-1	Sequence 1, Appli
26	823	52.4	11474	16	US-10-303-162-1	Sequence 1, Appli
27	823	52.4	11474	19	US-10-735-419-1	Sequence 1, Appli
28	823	52.4	11474	20	US-10-820-536-1	Sequence 1, Appli
29	823	52.4	11474	20	US-10-845-408-1	Sequence 1, Appli
30	823	52.4	11474	20	US-10-845-412-1	Sequence 1, Appli
31	823	52.4	11474	20	US-10-846-219-1	Sequence 1, Appli
32	823	52.4	11474	20	US-10-821-604-1	Sequence 1, Appli
33	823	52.4	11474	20	US-10-847-983-1	Sequence 1, Appli
34	823	52.4	11474	20	US-10-821-573-1	Sequence 1, Appli
35	823	52.4	11474	20	US-10-850-807-1	Sequence 1, Appli
36	823	52.4	11474	20	US-10-850-125-1	Sequence 1, Appli
37	823	52.4	11474	21	US-10-962-334-1	Sequence 1, Appli
38	823	52.4	11474	21	US-10-830-997-1	Sequence 1, Appli
39	823	52.4	11474	21	US-10-962-235-1	Sequence 1, Appli
40	823	52.4	876	9	US-09-816-028A-6	Sequence 6, Appli
41	817	52.0	876	15	US-10-303-161-6	Sequence 6, Appli
42	817	52.0	876	16	US-10-303-118-6	Sequence 6, Appli
43	817	52.0	876	16	US-10-303-128-6	Sequence 6, Appli
44	817	52.0	876	16	US-10-303-134-6	Sequence 6, Appli
45	817	52.0	876	16	US-10-303-162-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1

US-09-816-028A-2
; Sequence 2, Application US/09816028A
; Patent No. US20020042369A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/09/816,028A
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406


```

Db      835 AGATTACCTAGTGATATAAGCATATTATTTCAAAGGAAAA 873

RESULT 4
US-10-303-128-2
; Sequence 2, Application US/10303128
; Publication No. US20030157656A1
; GENERAL INFORMATION:
; APPLICANT: Wakarchuk, Michel
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,128
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-128-2

Alignment Scores:
Pred. No.: 1.57e-74 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 6

US-10-735-419-10 (1-294) x US-10-303-128-2 (1-876)

QY      8 LeuValCysGlyAenGlyProSerLeuLysAsnIleAspTyrIlyArgLeuProLysGln 27
Db      13 ATTTATTCGTGGAATGGACCAAGTTTAAAGAAATTTGATTTATTCAGAGACTACCAAAATGAT 72
QY      28 PheAspValPheArgCysAenGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGATTAGATGTAATCAATTTTATTTTGAGATAAATACTATCTTGGTAAATAA 132
QY      48 ValLysTyrValPhePheAenProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db      133 TGCAGGCGCATTTTACAACTCTTCTTTTGTGAAACAATACACTTTAAACAT 192
QY      68 LeuIleGluAenGluTyrAenIleGluAenIleValCysSerThrIleAenLeuGlu 87
Db      193 TTAATCCAAATCAAGATATGAGACCGCACTAATTTATGTTCTTAATTTACCAACCAAGCT 252
QY      88 TyrIleAspGlyPheGlnPheValAspAenPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAAAATGAAAAATTTTGAAAACTTTTACGATTTTCTGATGCTCATTTG 312
QY      108 GlyHisGluIleLysLysLeuLysAspPheAlaTyrIleLysTyrAenGluIle 127
Db      313 GGATATGATTTTTCACCAACATTAAGATTTTAAATGCTTTATTTTAAATTTTACGAAAT 372
QY      128 TyrAenArgGluIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAAATTTACCTCAGGGGTTTATATGTTGTCAGTAGCCCATAGCCCTAGGA 432

```

```

QY      148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAenAenLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCCGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
QY      168 PheAspAenAsnLysLysAenLeuAenLysCysThrGlyPheLysAenGlnLysPhe 187
Db      490 TTTGATACTAAACAAAAAATCTTTTAAATTTGGCTCCTTAATTTTAAATATGATAATTCA 549
QY      188 LysPheIleAenHisSerMetalaCysAspLeuGlnAlaLeuAspTyrIleuMetLysArg 207
Db      550 CACTATATCGACATAGTAAATAATACAGATATATAAAAGCTTTAGATTTCTAGAAAAAACT 609
QY      208 TyrAspValAenIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAATTAACACTATATTTGCTTATCTCTTAACAGCTCTTTTAGCAAAATTTTATAGAACTA 669
QY      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAenAsp 243
Db      670 GCGCCAAATTTAAATTTCAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
QY      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAAAAATTTTCAAATAATATT 774
QY      258 -----LysLysSerArgLeuLysGluAenLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAATAAATAAAGAAATATTTATTACAGTTGATAAAGATCTATTTA 834
QY      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACCTAGTGATATAAGCATTTATTTCAAAGGAAAA 873

```

RESULT 5

```

US-10-303-134-2
; Sequence 2, Application US/10303134
; Publication No. US20030157657A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/303,134
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-303-134-2

```

```

Alignment Scores:
Pred. No.: 1.57e-74 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18

```



```

Db      550 CACTATATCGGACATAGTAAAAATACAGATATATAAAAGCTTTAGAAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AapGluTyrPheLysLeu 223
Db      610 TACAAAATAAAACATATATGCTTATGTCCTTAACAGTCTTTTAGCAAAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCAAAATTTAAATTCAAATTTTATCATACAAGAAAA--AATAACTACACTAAAGAT 726
Qy      244 IleuIleProAspLysTyrAlaGlnGluArgTyrGly----- 257
Db      727 ATACTACATCTCTAGTAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAAATTAAGAAATATTATTACAGTTGATAAAGATCTATTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACCTAGTATAAAGCATTTATTTCAAAGGAAAA 873

RESULT 7
US-10-735-419-2
; Sequence 2, Application US/10735419
; Publication No. US20040180406A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/735,419
; CURRENT FILING DATE: 2003-12-11
; PRIOR APPLICATION NUMBER: US/09/816,028A
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
; OTHER INFORMATION: biosynthesis locus)
US-10-735-419-2

Alignment Scores:
Pred. No.: 1,57e-74 Length: 876
Score: 823.00 Matches: 161
Percent Similarity: 69.62% Conservative: 43
Best Local Similarity: 54.95% Mismatches: 71
Query Match: 52.39% Indels: 18
DB: 19 Gaps: 6

US-10-735-419-10 (1-294) x US-10-735-419-2 (1-876)

Qy      8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db      13 ATTATTGCTGGAATGGCAAGTTTAAAGAAATTTGATTATTCAAGACTACCAAAATGAT 72
Qy      28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTGATGATTTAGATGTAATCAATTTTATTGTAAGATTAATACTACTCTGTGTAATAAAA 132

```

```

Qy      48 ValIysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db      133 TGCAGGCGAGTATTTTACATCTCTATCTTTTTTTTGAACAATACTACACTTTTAAACAT 192
Qy      68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db      193 TTAATCCAAATCAAGAAATATGAGACCGCAACTAATTTATGTGTTCTTAATTACAAACCAAGCT 252
Qy      88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAAATGAAAATTTTGTAAAAACCTTTTACGATTATTTTCTGATGCTCATTTG 312
Qy      108 GlyHisGluIleIleLysLysLysAspPheAlaTyrIleLysTyrAsnGluIle 127
Db      313 GGTATATGATTTTTCACCAACTTAAGATTTAATGCTTTATTTTAAATTTACGAAAT 372
Qy      128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAAATTTACCTCAGGGTTTATATGTGTGAGTAGCCATAGCCCTTAGGA 432
Qy      148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
Qy      168 PheAspAsnAsnLysLysAsnLeuLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490 TTGTACTATAACAAAAAATCTTTTAAATTTGGCTCCTTAATTTTAAAAATGATAATTCA 549
Qy      188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGACATAGTAAAAATACAGATATAAAAGCTTTTAGAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAATAAAACTATATTTGCTTATGCTCTCAACAGCTTTTATGACAAATTTTATAGAACTA 669
Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTCAAATTTTATCATACAAGAAAA--AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db      727 ATACTCATACCTCTAGTAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAAATAAAATTAAGAAATATTATTACAGTTGATAAAGATCTATTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACCTAGTATATAAGCATTTATTTCAAAGGAAAA 873

RESULT 8
US-10-820-536-2
; Sequence 2, Application US/10820536
; Publication No. US20040203103A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/820,536
; CURRENT FILING DATE: 2004-04-07
; PRIOR APPLICATION NUMBER: 10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406

```



```
Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGCAAAATGAAATTTGTAAAAACTTTTACGATTATTTCTCGATGCTCATTTG 312
Qy 108 GlyHisGluIleIleLysLysLeuLysAspPheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGAATATGATTTTCAACAACTTAAAGATTTTAATGCTTATTTTAAATTTTCAGAAAT 372
Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAATTACTCAGGGTTTATATGTGTCAGTAGCCATAGCCCTAGGA 432
Qy 148 TyrLysSerIleTyrIleSerGlyLeAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT--TATGCT 489
Qy 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTTGATACTAAACAAATAATCTTTTAAATAATGGCTCCTCAATTTTAAATAATGATAATTCA 549
Qy 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGCATAGTAAATAACAGATATAAAGCTTTAGATTTCTAGAAAAACT 609
Qy 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAAACTATATTTCTTATGCTTATGCTTACAGTCTTTTACGAAATTTTATAGAACTA 669
Qy 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 GCGCCAAATTTAAATTTCAATTTATCATACAAAGAAAA--AATAACTACACTAAAGAT 726
Qy 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
Qy 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db 775 AATTTTAAAAATAAAATAAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTA 834
Qy 276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db 835 AGATTACCTAGTGATATAAGCAATTATTTCAAAGGAAAA 873
```

RESULT 10

```
US-10-845-412-2
; Sequence 2, Application US/10845412
; Publication No. US20040203113A1
; GENERAL INFORMATION:
; APPLICANT: Gilbert, Michel
; APPLICANT: Wakarchuk, Warren W.
; APPLICANT: National Research Council of Canada
; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
; FILE REFERENCE: 019633-000111US
; CURRENT APPLICATION NUMBER: US/10/845,412
; PRIOR FILING DATE: 2004-05-12
; PRIOR APPLICATION NUMBER: US/10/303,128
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US/09/816,028
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/118,213
; PRIOR FILING DATE: 1999-02-01
; PRIOR APPLICATION NUMBER: US 09/495,406
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 876
; TYPE: DNA
; ORGANISM: Campylobacter jejuni
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (1)..(876)
; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
; OTHER INFORMATION: biosynthesis locus)
US-10-845-412-2
```

Alignment Scores:

Pred. No.:	1-57e-74	Length:	876
Score:	823.00	Matches:	161
Percent Similarity:	69.62%	Conservative:	43
Best Local Similarity:	54.95%	Mismatches:	71
Query Match:	52.39%	Indels:	18
DB:	20	Gaps:	6

US-10-735-419-10 (1-294) x US-10-845-412-2 (1-876)

```
Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db 13 ATTATTGCTGGAATGGACCAAGTTTAAAGAAATTTGATTTTCAAGACTACCCAATGAT 72
Qy 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db 73 TTGTATGATTTTAGATGTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
Qy 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
Db 133 TGCNAGCCAGTATTTTCAATCTTCTTTTGAACATACTACTACTTTTAAAAACAT 192
Qy 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db 193 TTAATCCAAATCAAGAATATGAGACCGAACTAAATATGTTGTTCTTAATACCAACCAAGCT 252
Qy 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
Db 253 CATCTAGAAAAATGAAAAATTTGTAAAAAATTTTACGATTATTTTCTCATGCTCATTTG 312
Qy 108 GlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db 313 GGAATATGATTTTCAACAACTTAAAGATTTTAAATGCTTATTTTAAATTTCCAGAAAT 372
Qy 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db 373 TATTTCAATCAAGAATTTACTCTCAGGGTTTATATGTGTCAGTAGCCATAGCCCTAGGA 432
Qy 148 TyrLysSerIleTyrIleSerGlyLeAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db 433 TACAAAGAAATTTATCTTTTCGGGAATTTGATTTTATCAAAATGGGTCACTCT--TATGCT 489
Qy 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db 490 TTTTGATACTAAACAAATAATCTTTTAAATAATGGCTCCTCAATTTTAAATAATGATAATTCA 549
Qy 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db 550 CACTATATCGCATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAAACT 609
Qy 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db 610 TACAAATAAAACTATATTTGCTTATGCTTACAGTCTTTTACGAAATTTTATAGAACTA 669
Qy 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db 670 GCGCCAAATTTAAATTTCAATTTTATCATACAAAGAAAA--AATAACTACACTAAAGAT 726
Qy 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
Db 727 ATACTCATACTTCTAGTGAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
Qy 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db 775 AATTTTAAAAATAAAATAAAGAAAAATATTATTACAAAGTTGATAAAAGATCTATTA 834
```

us-10-735-419-10.rnpb

Thu Jun 16 13:06:59 2005

373 TATTTCAATCAAGAAATACCTCAGGGGTTTATATGTGTGCAGTAGCCATAGCCCTAGGA 432
 148 TrrLysSerIleTrrLysSerGlyIleAspPheTrrGlnAspThrAsnAsnLeuTrrAla 167
 433 TACAAGAAATTTATCTTCGGGAATTTGATTTTATCAAAATGGGTCTATCT---TATGCT 489
 168 PheAspAsnAsnLeuLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 490 TTTGATACCTAAACAAATAAATCTTTTAAATTTGGCTCCTTAATTTTAAATAATGATATTCA 549
 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTrrLeuMetLysArg 207
 550 CACTATATCGGACATAGTAAATAATACAGATATAAAAGCTTTAGAAATTTCTAGAAAACACT 609
 208 TrrAspValAsnIleTrrSerLeu-----AsnSer-----AspGluTrrPheLysLeu 223
 610 TACAAATAAAACTATATATGCTTATGCTCAACAGCTTTTAGCAAAATTTTATAGAACTA 669
 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTrrLysAsnAsp 243
 670 GCGCCAAATTTAAATTCAAAATTTTATCATACAGAAAAA---AATAACTACACTAAAGAT 726
 244 IleLeuIleProAspLysTrrAlaGlnGluArgTrrTrrGly----- 257
 727 ATACTCATACCTTCTAGTGAGCT-----TATGGAAAAATTTTCAAAAAATATT 774
 258 -----LysLysSerArgLeuLysGluAsnLeuHisTrrLysLeuLysLeuLysLeuLys 275
 775 AATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 834
 276 ArgLeuProSerAspIleLysHisTrrLysLysGluLys 288
 835 AGATTACCTAGTGTATATAAGCATTATTTCAAAGGAAAA 873

RESULT 12

US-10-821-604-2
 ; Sequence 2, Application US/10821604
 ; Publication No. US20040229263A1
 ; GENERAL INFORMATION:

APPLICANT: Gilbert, Michel
 APPLICANT: National Research Council of Canada
 TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 FILE REFERENCE: 019633-000111US
 CURRENT APPLICATION NUMBER: US/10/821.604
 CURRENT FILING DATE: 2004-04-08
 PRIOR APPLICATION NUMBER: 10/303,128
 PRIOR FILING DATE: 2002-11-21
 PRIOR APPLICATION NUMBER: US/09/816,028
 PRIOR FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: US 60/118,213
 PRIOR FILING DATE: 1999-02-01
 PRIOR APPLICATION NUMBER: US 09/495,406
 PRIOR FILING DATE: 2000-01-31
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 TYPE: DNA
 LENGTH: 876
 ORGANISM: Campylobacter jejuni
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(876)
 OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 OTHER INFORMATION: biosynthesis locus)
 US-10-821-604-2

Alignment Scores: 1.57e-74 Length: 876
 Pred. No.:

276 ArgLeuProSerAspIleLysHisTrrLysLysGluLys 288
 835 AGATTACCTAGTGTATATAAGCATTATTTCAAAGGAAAA 873

RESULT 11

US-10-846-219-2
 ; Sequence 2, Application US/10846219
 ; Publication No. US20040219638A1
 ; GENERAL INFORMATION:

APPLICANT: Gilbert, Michel
 APPLICANT: Wakarchuk, Warren W.
 TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 FILE REFERENCE: 019633-000111US
 CURRENT APPLICATION NUMBER: US/10/846.219
 CURRENT FILING DATE: 2004-05-14
 PRIOR APPLICATION NUMBER: US/09/816,028
 PRIOR FILING DATE: 2001-03-21
 PRIOR APPLICATION NUMBER: US 60/118,213
 PRIOR FILING DATE: 1999-02-01
 PRIOR APPLICATION NUMBER: US 09/495,406
 PRIOR FILING DATE: 2000-01-31
 NUMBER OF SEQ ID NOS: 49
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 TYPE: DNA
 LENGTH: 876
 ORGANISM: Campylobacter jejuni
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (1)..(876)
 OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 OTHER INFORMATION: biosynthesis locus)
 US-10-846-219-2

Alignment Scores: 1.57e-74 Length: 876
 Pred. No.: 823.00 Matches: 161
 Score: 69.62% Conservative: 43
 Percent Similarity: 54.95% Mismatches: 71
 Best Local Similarity: 52.39% Indels: 18
 Query Match: 20 Gaps: 6
 DB:

US-10-735-419-10 (1-294) x US-10-846-219-2 (1-876)

8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTrrLysArgLeuProLysGln 27
 13 ATATTGCTGGAAATGGACCAAGTTTAAAGAAATTTGATTTTCAGACTACCAATGAT 72
 28 PheAspValPheArgCysAsnGlnPheTrrPheGluAspArgTrrPheValGlyLysAsp 47
 73 TTTGATGATTTAGATGTTATCAATTTTATTTTGAAGATAAATACTATCTTGTTAAAAA 132
 48 ValLysTrrValPheAsnProPheValPheGluGlnTrrTrrSerLysLys 67
 133 TGCAAGCGCATTTTACAAATCTATTCTTTTTCGACCAATACACTACATTTAAACAT 192
 68 LeuIleGlnAsnGluTrrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 193 TTAATCCAAATCAAGATATGAGACCGAATTAATTTGTTTCTAATTAACAACCAAGCT 252
 88 TrrLysAspGlyPheGlnPheValAspAsnPheGluLeuTrrPheSerAspAlaPheLeu 107
 253 CATCTAGAAAATGAAAATTTTGTAAAAAATTTTACGATTATTTTCTGATGCTCATTTG 312
 108 GlyHisGluIleLysLysLysLysAspPheAlaTrrLysTrrAsnGluIle 127
 313 GGATATGATTTTTCACACCACTTAAAGATTTTAATGCTTATTTTAAATTTTCCAGAAAT 372
 128 TrrAsnArgGlnArgIleThrSerGlyValTrrMetCysAlaThrAlaValAlaLeuGly 147

Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservativity: 43
 Query Match: 54.95% Mismatches: 71
 DB: 52.39% Indels: 18
 Gaps: 6

US-10-735-419-10 (1-294) x US-10-821-604-2 (1-876)

QY 8 LeuValCysGlyAsnGlyProSerLeuLeuAsnLeuAspTyrLeuValGln 27
 Db 13 ATTTATGCTGGAATGACCAAGTTTAAAGAAATGATTTTCAAGACTACCAATGAT 72
 QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db 73 TTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
 QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
 Db 133 TGCAGGCGAGTATTTTCAATCTCTTTTGTGAACAATACTACACATTTAAACAT 192
 QY 68 LeuLeuGlnAsnGluGluTyrAsnLeuGluAsnLeuValCysSerThrLeuLeuGlu 87
 Db 193 TTAATCCAAATCAAGATATGAGCCGAATTAATATGTGTCAGTACCACTTAAACAT 252
 QY 88 TyrLeuAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db 253 CATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCTGATGCTCATTTG 312
 QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPhePheAlaTyrLeuLysTyrAsnGluLeu 127
 Db 313 GGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAT 372
 QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 Db 373 TATTTCAATCAAGATTTACCTCAGGGTTTATATGTGTCAGTACCACTAGCCCTAGGA 432
 QY 148 TyrLysSerIleTyrIleSerGlyLeuAspPheTyrGlnAspThrAsnLeuTyrAla 167
 Db 433 TACAAGAAATTTATCTTTCGGATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
 QY 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187

US-10-735-419-10 (1-294) x US-10-847-983-2 (1-876)

QY 8 LeuValCysGlyAsnGlyProSerLeuLeuAsnLeuAspTyrLeuValGln 27
 Db 13 ATTTATGCTGGAATGACCAAGTTTAAAGAAATGATTTTCAAGACTACCAATGAT 72
 QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db 73 TTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
 QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
 Db 133 TGCAGGCGAGTATTTTCAATCTCTTTTGTGAACAATACTACACATTTAAACAT 192
 QY 68 LeuLeuGlnAsnGluGluTyrAsnLeuGluAsnLeuValCysSerThrLeuLeuGlu 87
 Db 193 TTAATCCAAATCAAGATATGAGCCGAATTAATATGTGTCAGTACCACTTAAACAT 252
 QY 88 TyrLeuAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db 253 CATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCTGATGCTCATTTG 312
 QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPhePheAlaTyrLeuLysTyrAsnGluLeu 127
 Db 313 GGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAT 372
 QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 Db 373 TATTTCAATCAAGATTTACCTCAGGGTTTATATGTGTCAGTACCACTAGCCCTAGGA 432
 QY 148 TyrLysSerIleTyrIleSerGlyLeuAspPheTyrGlnAspThrAsnLeuTyrAla 167
 Db 433 TACAAGAAATTTATCTTTCGGATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
 QY 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187

Alignment Scores:
 Pred. No.: 1.57e-74 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservativity: 43
 Query Match: 54.95% Mismatches: 71
 DB: 52.39% Indels: 18
 Gaps: 6

US-10-735-419-10 (1-294) x US-10-847-983-2 (1-876)

QY 8 LeuValCysGlyAsnGlyProSerLeuLeuAsnLeuAspTyrLeuValGln 27
 Db 13 ATTTATGCTGGAATGACCAAGTTTAAAGAAATGATTTTCAAGACTACCAATGAT 72
 QY 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 Db 73 TTTGATGATTTAGATGATTAATCAATTTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
 QY 48 ValLysTyrValPhePheAsnProPheValPheGluGlnTyrTyrThrSerLysLys 67
 Db 133 TGCAGGCGAGTATTTTCAATCTCTTTTGTGAACAATACTACACATTTAAACAT 192
 QY 68 LeuLeuGlnAsnGluGluTyrAsnLeuGluAsnLeuValCysSerThrLeuLeuGlu 87
 Db 193 TTAATCCAAATCAAGATATGAGCCGAATTAATATGTGTCAGTACCACTTAAACAT 252
 QY 88 TyrLeuAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 Db 253 CATCTAGAAAATGAAAATTTTGTAAAACCTTTTACGATTATTTTCTGATGCTCATTTG 312
 QY 108 GlyHisGluLeuLeuLysLysLeuLysAspPhePheAlaTyrLeuLysTyrAsnGluLeu 127
 Db 313 GGATATGATTTTTCACCAACTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAT 372
 QY 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 Db 373 TATTTCAATCAAGATTTACCTCAGGGTTTATATGTGTCAGTACCACTAGCCCTAGGA 432
 QY 148 TyrLysSerIleTyrIleSerGlyLeuAspPheTyrGlnAspThrAsnLeuTyrAla 167
 Db 433 TACAAGAAATTTATCTTTCGGATTTGATTTTATCAAAATGGGTCACTCT---TATGCT 489
 QY 168 PheAspAsnAsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187

US-10-847-983-2

Sequence 2, Application US/10847983
 Publication No. US20040229272A1
 GENERAL INFORMATION:

us-10-735-419-10.rnpb

Thu Jun 16 13:06:59 2005

490 TTTGATATAAACAACAAAAATCTTTTAAATGGCTCTCTAAATTTTAAAAATGATAAATTC 549
 188 LysPheLeuAenHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 550 CACTATATCGGCACATAGTAAAAATACAGATATAAAGCTTTAGAAATTTCTAGAAAAACT 609
 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 610 TACAAAATAAACAATATATGCTTATGCTCTAAAGCTTTTACAGCAATTTTATAGAACTA 669
 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
 670 GCGCCAAATTAATTAATTTTATCATCAAGAAAAA---AATAACTACACTAAAGAT 726
 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuLeu 275
 775 AATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 834
 276 ArgLeuProSerAspIleLysHisTyrLysLeuLysGluLys 288
 835 AGATTACCTAGTATATAAAGCATTATTTCAGAGGAAAA 873

RESULT 14
 US-10-821-573-2
 ; Sequence 2, Application US/10821573
 ; Publication No. US2004022931A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: Wakarchuk, Warren W.
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/821,573
 ; PRIOR FILING DATE: 2004-04-08
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR FILING DATE: 2001-03-21
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 876
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(876)
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS)
 ; OTHER INFORMATION: biosynthesis locus)
 ; US-10-821-573-2

Alignment Scores:
 Pred. No.: 1,57e-74 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Best Local Similarity: 54.95% Mismatches: 71
 Query Match: 52.39% Indels: 18
 DB: 20 Gaps: 6

US-10-735-419-10 (1-294) x US-10-821-573-2 (1-876)
 Qy 8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27

13 ATTATGCTGGAATGGACCAAGTTTAAAAAGAAATTTGATTATTATCAAGACTACCAATGAT 72
 28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
 73 TTTGATGATTATGATTAATCAATTTTATTTTGAAGATAAATACTACTCTTGGTAAAAA 132
 48 VallysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
 133 TGCAAGGAGTATTTTACAACTCTTCTTTTGTGAAACAATACTACACTTTTAAACAT 192
 68 LeuIleGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
 193 TTAATCCAAAATCAAGATATGAGACCGAACTAATATATGTGTCTTAATTACACCAAGCT 252
 88 TyrIleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeu 107
 253 CATCTAGAAAATGMAAATTTTGTAAAACTTTTTCAGATATTATTTCTGATGCTCATTTG 312
 108 GlyHisGluIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
 313 GGATATGATTTTTCACCAACTTTAAAGATTTTAAATGCTTATTTTAAATTTACGAAAT 372
 128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
 373 TATTTCAATCAAGAAATTTACTCAGGGGTTTATATGTGTGAGTAGCATAGCCCTAGGA 432
 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
 433 TACAAAGAAATTTTATCTTCGGGAATTTGATTTTATCAAAATGGTCACTCT--TATGCT 489
 168 PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPhePheLysAsnGlnLysPhe 187
 490 TTTGATACTAAACAAAAAATCTTTTAAATTTGGCTCTCTAATTTTAAAAATGATTAATCA 549
 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 550 CACTATATCGGCACATAGTAAAAATAAATAAAGCTTTTAAATTTTCTAGAAAAACT 609
 208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
 610 TACAAAATAAACAATATATGCTTATGCTCTTAAAGCTTTTACCAAAATTTTATAGAACTA 669
 224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
 670 GCGCCAAATTTAAATTTTAAATTTTATCATCAAGAAAAA---AATAACTACACTAAAGAT 726
 244 IleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGly----- 257
 727 ATACTCATACCTTCTAGTGAGGCT-----TATGGAATAATTTTCAAAAAATATT 774
 258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuLysAspLeuLeu 275
 775 AATTTTAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 873
 835 AGATTACCTAGTATATAAAGCATTATTTCAGAGGAAAA 873

RESULT 15
 US-10-850-807-2
 ; Sequence 2, Application US/10850807
 ; Publication No. US20040259140A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gilbert, Michel
 ; APPLICANT: Wakarchuk, Warren W.
 ; APPLICANT: National Research Council of Canada
 ; TITLE OF INVENTION: Campylobacter Glycosyltransferases for Biosynthesis of
 ; TITLE OF INVENTION: Gangliosides and Ganglioside Mimics
 ; FILE REFERENCE: 019633-000111US
 ; CURRENT APPLICATION NUMBER: US/10/850,807
 ; CURRENT FILING DATE: 2004-05-21
 ; PRIOR APPLICATION NUMBER: US/09/816,028

; PRIOR FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: US 60/118,213
 ; PRIOR FILING DATE: 1999-02-01
 ; PRIOR APPLICATION NUMBER: US 09/495,406
 ; PRIOR FILING DATE: 2000-01-31
 ; NUMBER OF SEQ ID NOS: 49
 ; SOFTWARE: Patentin Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 876
 ; TYPE: DNA
 ; ORGANISM: Campylobacter jejuni
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(876)
 ; OTHER INFORMATION: bifunctional alpha-2,3/alpha 2,8-sialyltransferase
 ; OTHER INFORMATION: Campylobacter sialyltransferase II (cstII) from C. jejuni
 ; OTHER INFORMATION: strain OH4384 (ORF 7a of lipooligosaccharide (LOS))
 US-10-850-807-2

Alignment Scores:

Pred. No.: 1.57e-74 Length: 876
 Score: 823.00 Matches: 161
 Percent Similarity: 69.62% Conservative: 43
 Best Local Similarity: 54.95% Mismatches: 71
 Query Match: 52.39% Indels: 18
 DB: 20 Gaps: 6

US-10-735-419-10 (1-294) x US-10-850-807-2 (1-876)

```

Qy      8 LeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGln 27
Db      13 ATTTTGTCTGGAATTTGACCAAGTTTAAAGAAATTTGATTATTCAGACTACCAATGAT 72
Qy      28 PheAspValPheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAsp 47
Db      73 TTTGATGATTATGATGATTAATCAATTTATTTGAAGATAAATACTATCTTGGTAAAAA 132
Qy      48 ValLysTyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLys 67
Db      133 TGCAAGCGCATTTTACAAATCTCTTTTGTGAACTAATACTATGTTCTTAATTACCAACCAAGCT 252
Qy      68 LeuLeuGlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGlu 87
Db      193 TTAATCCAAATCAAGATATGAGCCGAACTAATAATGTTCTTAATTACCAACCAAGCT 252
Qy      88 TyrIleAspGlyPheGlnPheValAspAsnGluLeuTyrPheSerAspAlaPheLeu 107
Db      253 CATCTAGAAAATGAAAATTTGTAAACTTTTACGATTATTTTCTGATGCTCATTTG 312
Qy      108 GlyHisGluIleIleLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIle 127
Db      313 GGATATGATTTTTCAAACCACTTAAAGATTTTAATGCTTATTTTAAATTTTCAGAAAT 372
Qy      128 TyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGly 147
Db      373 TATTTCAATCAAGAAATTAACCTCAGGGTTTATATGTGTGAGTACCCATAGCCCTAGGA 432
Qy      148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
Db      433 TACAAAGAAATTTATCTTTCCGGGAATGATTTTATCAAAATGGGTCATCT---TATGCT 489
Qy      168 PheAspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
Db      490 TTTGATACTAAACAAAATAATCTTTTAAATTTGGCTCCTAATTTTAAATGATAATTC 549
Qy      188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
Db      550 CACTATATCGGACATAGTAAATAACAGATATAAAGCTTTAGAAATTTCTAGAAAAAACT 609
Qy      208 TyrAspValAsnIleTyrSerLeu-----AsnSer-----AspGluTyrPheLysLeu 223
Db      610 TACAAAATAAACTATATGCTTATCTCTACAGTCTTTTAGCAAAATTTTATAGAACTA 669
  
```

```

Qy      224 AlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAsp 243
Db      670 GCGCCAAATTTAAATTCAAATTTATCATACAAGAAAA---AATAACTACACTAAAGAT 726
Qy      244 IleLeuIleProAspLysTyrAlaGlnGluAtgTyrTyGly-----TATGAAAAATTTTCAAAAAATATT 774
Db      727 ATACTCATACCTTCTAGTCAGGCT-----TATGAAAAATTTTCAAAAAATATT 774
Qy      258 -----LysLysSerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIle 275
Db      775 AATTTTAAAAATAAAAAATTAAGAAAAATATTATTACAAAGTTGATATAAAGATCTATTA 834
Qy      276 ArgLeuProSerAspIleLysHisTyrLeuLysGluLys 288
Db      835 AGATTACTAGTATATAAAGCAATTTTCAAGAAAA 873
  
```

Search completed: June 14, 2005, 22:46:46
 Job time : 678 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 14, 2005, 18:10:40 ; Search time 3688 Seconds
(without alignments)
3034.408 Million cell updates/sec

Title: US-10-735-419-10
Perfect score: 1571
Sequence: 1 MSMMINALVCGNGSLKNID.....IRLPDIKHLYLKEKYANKNR 294

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_epool/US10735419/runat_13062005_165725_12664/app_query.fasta_1.455
-DB=EST -OPWT=fastap -SUFFIX=rst -MINMATCH=0.1 -ILOOPCL=0 -ILOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=spt -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10735419 @CGN 1 1 4352 @runat_13062005_165725_12664 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	118	7.5	828 7 CN594379 TTE000141
2	117.5	7.5	981 9 CN507888 AL437553 T7 end of
3	116	7.4	515 5 BQ577290 PFESToab1
4	113	7.2	574 5 BQ497975
5	113	7.2	894 7 CK573341
6	111.5	7.1	774 4 BM167106
7	107	6.8	1123 8 A2935754
8	106.5	6.8	666 6 CB855294
9	106.5	6.8	678 4 BU371051

C 11	104.5	6.7	833	4	BJ365708
C 12	104.5	6.7	924	8	AZ679344 ENTME86TR
C 13	104	6.6	973	8	AZ546559 ENTDR06TF
C 14	103	6.6	699	4	AZ684114 ENTQ74TF
C 15	103	6.6	787	4	BJ344875 BJ344875
C 16	103	6.6	954	9	BM162809 EST565332
C 17	102.5	6.5	915	8	CL049993 CH216-71H
C 18	102.5	6.5	922	8	AZ531151 ENTBN36TF
C 19	102	6.5	645	1	AU038165 ENTDRK95TF
C 20	102	6.5	1001	9	AU038165 AU038165
C 21	101.5	6.5	498	5	CNS07C9G T7 end of
C 22	101.5	6.5	795	8	BQ596294 PFESToab3
C 23	101.5	6.5	825	4	BZ396831 E1NAD41TR
C 24	101	6.4	1153	8	BJ371205 BJ371205
C 25	100.5	6.4	715	4	CC270116 CC270116
C 26	100.5	6.4	745	4	BJ392327 BJ392327
C 27	100.5	6.4	753	8	BJ365777 BJ365777
C 28	100.5	6.4	829	4	BH590061 BQJZ45TR
C 29	100.5	6.4	834	4	BJ336376 BJ336376
C 30	100.5	6.4	952	8	BJ335751 BJ335751
C 31	100	6.4	716	8	BH161189 ENTQV23TR
C 32	100	6.4	722	4	BH441159 BOGYH37TF
C 33	99.5	6.3	540	4	BI814723 BJ333627
C 34	99.5	6.3	706	4	BJ368813 BJ368813
C 35	99.5	6.3	984	9	CL867919 Gma_01_11
C 36	99	6.3	604	8	CL867919 B78634
C 37	99	6.3	630	7	CF788898 B78634
C 38	99	6.3	646	8	AZ525450 CP00054A Cp
C 39	99	6.3	655	8	AF252603 251PbC05
C 40	99	6.3	762	8	AZ183508 SP_1001_A
C 41	99	6.3	1092	8	BZ548462 pacsl-60
C 42	99	6.3	1487	5	BP937769 BP937769
C 43	99	6.3	2167	3	BC021319 Mus muscu
C 44	98.5	6.3	482	4	BG603768 EST502858
C 45	98.5	6.3	828	6	CA457324 AGENCOURT

ALIGNMENTS

RESULT 1
CN594379
LOCUS CN594379 828 bp mRNA linear EST 31-AUG-2004
DEFINITION TTE000141.05 Normalized large Tetrahymena thermophila cDNA, mRNA
sequence.
ACCESSION CN594379
VERSION CN594379.1 GI:47046181
KEYWORDS EST.
SOURCE Tetrahymena thermophila
ORGANISM Tetrahymena thermophila
REFERENCE 1 (bases 1 to 828)
AUTHORS Garg J., Pearlman R.E. and Carlton J.
TITLE PFdbPub (http://amcebida.bcm.umontreal.ca/public/pepdb/agrm.php)
JOURNAL Tetrahymena thermophila (TIGR)
COMMENT Unpublished (2004)
Contact: PFdb
Deparment de Biochimie, Universite de Montreal
Email: pepdb-curator@bch.umontreal.ca
Plate: 1398.
FEATURES
source
1..828
/organism="Tetrahymena thermophila"
/mol_type="mRNA"
/db_xref="taxon:5911"
/clone_lib="Normalized large"
ORIGIN
Alignment Scores: 0.00278 Length: 828
Pred. No.: 118.00 Matches: 68
Percent Similarity: 39.92%

```
Best Local Similarity: 26.36% Mismatches: 79
Query Match: 7.51% Indels: 76
DB: 7 Gaps: 15

US-10-735-419-10 (1-294) x CNS594379 (1-828)

Qy 72 GluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGluTyrIleAspGly 91
   |||||
   |||||
   |||||
Db 1 GAAGAATATAAATCTATAAGTATTAT-----27

Qy 92 PheGlnPheValAspAspPheGluLeuTyrPheSerAspAlaPheLeuGluHisGluIle 111
   |||||
   |||||
   |||||
Db 28 TTTAAATTTTAAAAAATTTTAAAAATATTATTT-----ATTTTAAATTTTAAAAAT 78
   |||||
   |||||
   |||||

Qy 112 IleIleValys-----LeuLysAspPhePheAla 120
   |||||
   |||||
   |||||
Db 79 ATTAAAAAAGATGATATTTTATTATAAATGAATATATTATTATAAATTAACTTATTATTT 138
   |||||
   |||||
   |||||

Qy 121 TyrIleLys-----TyrAsnGluIleTyrAsn 129
   |||||
   |||||
   |||||
Db 139 TATATAAAGATGATTTCTAGAAATAAAATTAACCTTAATTTTATATAAT--ATTATAAT 195
   |||||
   |||||
   |||||

Qy 130 ArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGluTyrLys 149
   |||||
   |||||
   |||||
Db 196 AATAATTATATAAATAAATATATATATATATTAACATATAATACAAAATTTAAATATAAT 255
   |||||
   |||||
   |||||

Qy 150 SerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAlaPheAsp 169
   |||||
   |||||
   |||||
Db 256 AATCTA-----ATTAATATAAATAAATAAATAAATAAATAAATAAATAAATAA 297
   |||||
   |||||
   |||||

Qy 170 Asn-----AsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsn 184
   |||||
   |||||
   |||||
Db 298 AATAATTATAGTAAATATAATAATAATAATAATAATAATAATAATAATAATAAATAA 357
   |||||
   |||||
   |||||

Qy 185 GlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeu 204
   |||||
   |||||
   |||||
Db 358 TTTAAATAAAT--ATAAATTTATAATAATA-----TATATA 393
   |||||
   |||||
   |||||

Qy 205 MetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyrPheLysLeuAla 224
   |||||
   |||||
   |||||
Db 394 TATAAATTTAAAAAATTTTATAAATTAATTAATAATAATAAAT--TGAAAGAAATYATA 450
   |||||
   |||||
   |||||

Qy 225 ProAspIleGlySerAspPheValLeuSerLysLysProLysLys-----239
   |||||
   |||||
   |||||
Db 451 TTGAATTTAAAAATTTTATATATAAATAGTAAAGACCACTTAAACAACTTTAAATATGAT 510
   |||||
   |||||
   |||||

Qy 240 -----TyrIleAsnAspIleIleuleProAspLysTyrAlaGlnGluArgTyrTyr 256
   |||||
   |||||
   |||||
Db 511 AAAAGTTTATATATATAAATAATTTTAGATGGATTAGAATATGCTATG-----TTAACA 564
   |||||
   |||||
   |||||

Qy 257 GlyLysLysSerArgLeuLysGluAsnLeu-----HisTyrLysLeuIleLysAsp 273
   |||||
   |||||
   |||||
Db 565 AATAAANAATCATATAAANAATAAANAATAAATAAATAAATAAATAAATAAATAAATAA 621
   |||||
   |||||
   |||||

Qy 274 LeuIleArgLeuProSerAspIleLysHisTyrLysLysGluLysTyrAlaAsn 291
   |||||
   |||||
   |||||
Db 622 TTTAAAAAATACCAAGTAAATA-----GAAAAATATATAAAT 660
   |||||
   |||||
   |||||

RESULT 2
CNS07B8B/c 981 bp DNA linear GSS 08-JUL-2001
LOCUS T7 end of clone BC0AA010A05 of library BC0AA of strain CBS 767 of
DEFINITION Debaryomyces hansenii, genomic survey sequence.
ACCESSION AL437553
VERSION AL437553.1 GI:12220966
KEYWORDS GSS.
SOURCE Debaryomyces hansenii (anamorph: Candida famata)
ORGANISM Debaryomyces hansenii
          Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
          Saccharomycetales; Saccharomycetaceae; Debaryomyces.
REFERENCE 1 (bases 1 to 981)
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
          Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,
          de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,
          Malpertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
          Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
          Wincker,P. and Weissenbach,J.
          Genomic exploration of the hemiascomycetous yeasts: 1. A set of
          yeast species for molecular evolution studies
          FEBS Lett. 487 (1), 3-12 (2000)
          20584711
          11152876
          2 (bases 1 to 981)
          Lepingle,A., Casaregola,S., Neuveglise,C., Bon,E., Nguyen,H.,
          Artiguenave,F., Wincker,P. and Gaillardin,C.
          Genomic exploration of the hemiascomycetous yeasts: 14.
          Debaryomyces hansenii var. hansenii
          FEBS Lett. 487 (1), 82-86 (2000)
          20584724
          11152889
          3 (bases 1 to 981)
          Genoscope.
          Direct Submission
          Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
          2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex - FRANCE (E-mail :
          seque@genoscope.cns.fr - Web : www.genoscope.cns.fr)
          This GSS is part of a random genomic sequencing program of thirteen
          yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
          exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,
          lactic var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
          angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
          Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
          5 kb were prepared and both extremities were sequenced. See
          keywords for description of this sequence and for the sequence of
          the other extremity of this insert.
          Location/Qualifiers
          1..981
            /organism="Debaryomyces hansenii"
            /mol_type="genomic DNA"
            /strain="CBS 767"
            /variety="hansenii"
            /db_xref="taxon:4959"
            /clone="BC0AA010A05"
            /clone_lib="BC0AA"
            /note="end : T7"

FEATURES
ORIGIN
Alignment Scores:
Pred. No.: 0.00398 Length: 981
Score: 117.50 Matches: 73
Percent Similarity: 35.67% Conservative: 49
Best Local Similarity: 21.35% Mismatches: 109
Query Match: 7.48% Indels: 111
DB: 9 Gaps: 16

US-10-735-419-10 (1-294) x CNS07B8B (1-981)

Qy 5 IleAsnAlaLeuValCysGlyAsnGlyProSerLeuLysAsnIleAspTyrLysArgLeu 24
   |||||
   |||||
   |||||
Db 936 CTTGAAAATTTATGTTTAATTAATGATCCAACGATCCAG-----GACTATGGAAGAATA 883
   |||||
   |||||
   |||||

Qy 25 ProLysGln-----PheAspValPheArg 32
   |||||
   |||||
   |||||
Db 882 AAGATGTTCTTCAGTATATTGTCAAAATAATAATCTCCAGATTTTGGTTTCCCGTTTTTAGA 823
   |||||
   |||||
   |||||

Qy 33 -----CysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAspValLysTyr 50
   |||||
   |||||
   |||||
Db 822 GAATTATTATCTCAATCATTTATTCATATAAACAACCTGGGATACAGAA-----775
   |||||
   |||||
   |||||

Qy 51 ValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLysLeuIleGln 70
   |||||
   |||||
   |||||
Db 774 -----TATTATCCAGAGTATGTTTCTTTAAACTATATACGTATAAATATATATTCGTGAT 721
   |||||
   |||||
   |||||

Qy 71 AsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGluTyrIleAsp 90
   |||||
   |||||
   |||||
```

Db 720 AATGAGAGGATTATTTAGTGAATTTTATTAACCTTATTAATTC---TATATTCAT 664

Qy 91 GlyPheGln-----PheValAspAsnPheGluLeuTyThrSerAspAlaPhe---Leu 107

Db 663 GCAATGGATAATCCCTTTTATGACGCGTAATGATATTTCTGCAACTGCTACTTATGGAAC 604

Qy 108 GlyHisGluIleIleLysLeuLysAspPhe----- 118

Db 603 GGTCAAGAAATGATTTGGAATAATGATTTTATAAAATTTGGACTATATCGAATA 544

Qy 118 ----- 118

Db 543 ATCTTAATTTATAAAGACTTAAGACTAATATAAAGAACTCGAAGCATTATTATTCATGAA 484

Qy 119 -----PheAlaTyrlleLysTy-----AsnGluIleTyThrAsnArgGlnArgIle 133

Db 483 TTAACATTTGCTTATCTTCTGAGTTTCCAAACCAACATAAATCTGCTAAATCTAGATCAAAT 424

Qy 134 ThrSerGlyValTyThrMetCysAlaThrAlaValAlaLeuGlyTyThrLysSerIleTyrlle 153

Db 423 -----ATTGTAATTTTGAAGAAATACAAAGGTTTAT 391

Qy 154 SerGlyIleAspPheTyThrAsnAsnLeuTyAla-----PheAspAsnAsn 171

Db 390 AAACCCAAATATTTTCAAGAGAAACTGATAAT---TATTCAAACAGATTTCGATAATAT 334

Qy 172 LysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPheLysPheIleAsn 191

Db 333 AAAAAGAAGCAACTGAAG-----AAAAATAATAATATGAACATAGGAAT 289

Qy 192 HisSerMetAlaCysAspLeuGlnAlaLeuAspTyThrLeuMetLysArgTyAspValAsn 211

Db 288 CAAATTTCTTCAAGAGAT----- 271

Qy 212 IleTyThrLeuAsnSerAspGluTyThrPheLysLeuAlaProAspIleGlySerAspPhe 231

Db 270 -----AATGATGAAATTAACAA-----ATTGTTAGTCACAT 238

Qy 232 ValLeuSerLysLysProLysTyThrIleAsnAspIleLeuIleProAspLysTyAla 251

Db 237 ACTGATATAAAATATCTAGTTTTCATATAATAATATATTTTCAAGTGACCAACCAATC 178

Qy 252 GlnGluArgTyThrGlyLysLysSerArgLeuLysGluAsnLeuHisTyLysLeu--- 270

Db 177 AATAATCATCATCTATGATAATAATAAATAGATAGATAGCAAAATATGTTTTCGAC 118

Qy 271 -----IleLysAspLeuIleArgLeuProSerAsp 280

Db 117 AGCGGAGCAAGGTTACTGTGTAATGATATATAAACAATTAATAATAAAGGAAT 58

Qy 281 IleLys 282

Db 57 TTAGAA 52

RESULT 3

BO577290

LOCUS PFESTOab16h12.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium falciparum 3D7 cDNA 5', mRNA sequence.

DEFINITION

ACCESSION BO577290

VERSION BO577290.1

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 515)

Eukaryota; Alveolata; Apicomplexa; Haemosporidia; Plasmodium.

Tang, K., Cole, R., Chakrabarti, D., Haywood, R., Clifton, S., Pape, D., Marra, M., Hillier, L., Martin, J., Wyllie, T., Dente, M., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Jentes, E., Ronko, I., Tsagaris, V., Richey, J., Belaygorod, L., Franklin, C., Carr, L., Grow, A., Maguire, L., Richey, J., Wadkins, J., Kennedy, S., Levinso, D., Waterston, R., Wilson, R. and Sibley, D.

WashU Plasmodium EST Project

TITLE

JOURNAL COMMENT

Unpublished (2001)

Contact: L. David Sibley

WashU Plasmodium EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

Library was constructed by Debopam Chakrabarti DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: L. David Sibley (sibley@borcim.wustl.edu), Washington University

Seq primer: -40UP from Gibco

High quality sequence stop: 426.

FEATURES

source

Location/Qualifiers

1..515

/organism="Plasmodium falciparum 3D7"

/mol_type="mRNA"

/db_xref="taxon:36329"

/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"

/clone_lib="Plasmodium falciparum 3D7 asexual cDNA"

/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2: XhoI; Library was constructed by Debopam Chakrabarti. Total RNA samples were isolated from mixed stage saponin(0.1%) lysed P. falciparum 3D7 infected erythrocytes by the acidic guanidinium-phenol chloroform method. The poly A+ RNA was isolated by the polyAT-Tract Magna isolation system (Promega, WI) using streptavidin RNA sphere particles. Directional cDNA libraries were constructed by oligo d(T) priming of poly(A)+ RNA (5mg) into EcoRI and XhoI sites of 1 ZapII vector using the Zap cDNA synthesis kit (Stratagene, CA). The average size of the cDNA inserts in the library was between 1.0 and 1.5kb. Clones were mass excised using the ExAssist helper phage (Stratagene), the phagemids were precipitated with PEG 8000 and extracted with phenol/chloroform. Phagemid DNA was electroporated into DH10B cells."

ORIGIN

Alignment Scores:

Pred. No.: 0.00258 Length: 515

Score: 116.00 Matches: 47

Percent Similarity: 42.11% Conservatives: 41

Best Local Similarity: 22.49% Mismatches: 63

Query Match: 7.38% Indels: 58

DB: 5 Gaps: 11

US-10-735-419-10 (1-294) x BO577290 (1-515)

Qy 56 PheValPhePheGluGlnTyThrSerLysLysLeuIleGlnAsnGluGluTyThrAsn 75

Db 1 TTTTCTTTTATAAACAATATATTTATATGAG-----CTAAGGAAGGAATATATAT 51

Qy 76 IleGluAsnIleValCysSerThrIleAsnLeuGluTyThrIleAspGlyPheGlnPheVal 95

Db 52 AATCAAAATATATATGTCG-----AATAAAATGGAACATCAAACTTTATATA 105

Qy 96 AspAsnPheGluLeuTyThrPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLeu 115

Db 106 CCTTTTAAGAAGACATATATCGAA-----GATTATGAACCTTTAATGAAGA 156

Qy 116 LysAspPhePheAlaTyrlleLysTyThrAsnGluIleTyThrAsnArgGlnArgIleThrSer 135

Db 157 CGAAATATTTTATAGAGCGTATCTATCATCATGATCTGATATATAAT----- 201

Qy 136 GlyValTyThrMetCysAlaThrAlaValAlaLeuGlyTyThrLysSerIleTyrlleSerGly 155

Db 202 -----AAAAATTTATTATTCCTCGA 222

Qy 156 -----IleAspPheTyThrGlnAspThr-----AsnAsnLeu 165

Db 223 GATAACACCAGTATTTATGATAATACAGATGATGATATATAATAAGATGATATATCAT 282

Qy 166 TyrAlaPheAspAsnAsnLysLysAsnLeu-----LeuAsn 177
 Db 263 TATATTTTAAATAATAAATAATATAGTGAATATGTTATGATAAAGATACACAT 342
 Qy 178 LysCysThrGlyPheLysLysGlnLysPheLysPheLeuAsnHisSerMetAlaCysAsp 197
 Db 343 ACACATACCAATATATATCTGATATATACAAAAAATAACGATACACATAATGAT 402
 Qy 198 LeuGlnAlaLeuAspTyrLeuMetLysArgTyrAspValAsnLleTyrSerLeuAsnSer 217
 Db 403 TTTCGTACATG-----TTAATAAATAATACATAATAATATATATAGTCTTAACAAA 456
 Qy 218 Asp-----GluTyrPheLysLeuAlaProAspLleGlySerAspPheValLeuSerLys 235
 Db 457 GATTACACAATTTAT-----AATGATGAATATTA----- 486
 Qy 236 LysProLysLysTyrIleAsnAspIle 244
 Db 487 AAAGTTGAAGAAATAAATAATGAATA 513
 RESULT 4
 BU497975
 LOCUS PfESToab91a03.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium
 DEFINITION falciparum 3D7 cDNA 5', mRNA sequence.
 ACCESSION BU497975
 VERSION BU497975.1 GI:22794169
 KEYWORDS EST.
 SOURCE Plasmodium falciparum 3D7
 ORGANISM Plasmodium falciparum 3D7
 1 (bases 1 to 574)
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 AUTHORS Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
 Maiza,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B.,
 Bowers,I., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
 Tsagareishvili,R., Belyagorod,L., Franklin,C., Carr,L., Grow,A.,
 Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
 Waterston,R., Wilson,R. and Sibley,D.
 WASHU Plasmodium EST Project
 Unpublished (2001)
 CONTACT: L. David Sibley
 WASHU Plasmodium EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library was constructed by Debopam Chakrabarti DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: L. David Sibley
 (sibley@borcim.wustl.edu), Washington University
 Seq primer: -40UP from Gibco
 High quality sequence stop: 421.
 Location/Qualifiers
 1..574
 /organism="Plasmodium falciparum 3D7"
 /mol_type="mRNA"
 /db_xref="taxon:36329"
 /lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
 /clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
 /note="Vector: pBluescript SK plus; Site 1: EcoRI; Site_2:
 XhoI; library was constructed by Debopam Chakrabarti.
 Total library samples were isolated from mixed stage
 saponin(0.1%) -lysed P. falciparum 3D7 infected
 erythrocytes by the acidic guanidium-phenol chloroform
 method. The poly A+ RNA was isolated by the polyAT-Tract
 mRNA isolation system (Promega, WI) using streptavidin
 Magnasphere particles. Directional cDNA libraries were
 constructed by oligo d(T) priming of poly(A)+ RNA (5mg)
 into EcoRI and XhoI sites of 1 ZapII vector using the Zap
 cDNA synthesis kit (Stratagene, CA). The average size of the
 cDNA inserts in the library was between 1.0 and 1.5kb.
 Clones were mass excised using the ExAssist helper phage

(Stratagene), the phagemids were precitptated with PEG
 8000 and extracted with phenol/chloroform. Phagemid DNA
 was electroporated into DH10B cells."

ORIGIN
 Alignment Scores:
 Pred. No.: 0.00678 Length: 574
 Score: 113.00 Matches: 52
 Percent Similarity: 40.64% Conservative: 37
 Best Local Similarity: 23.74% Mismatches: 76
 Query Match: 7.19% Indels: 54
 DB: 5 Gaps: 11
 US-10-735-419-10 (1-294) x BU497975 (1-574)
 Qy 43 PheValGlyLysAspValLysTyrVal-----PhePheAsnPropheValPhePhe 59
 Db 54 TATATAGATAAAGACACACATTTTATTATTCAAACATTTTATAATAATTTTATTATTAT 113
 Qy 60 GluGlnTyrTyrThrSerLysLysLeuIleGlnAsnGlu-----GluTyrAsnIleGlu 77
 Db 114 ATTGAATAATTATGATAAATATGATGATCAATGCAATGACACAGGCAAGCAATATAA 173
 Qy 78 AsnIleValCysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsn 97
 Db 174 AATATAAAAAATAATAATACGAATGGACAGAAAGGAT-----ATTATAAAT 221
 Qy 98 PheGluLeuTyrPheSerAspAlaPheLeuGlyHisGluIleLleLysLysLeuLysAsp 117
 Db 222 GATGAATATTTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 281
 Qy 118 PhePhe-----AlaTyrIleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSer 135
 Db 282 TCITTCATGAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 338
 Qy 136 GlyValTyrMetCysAlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGly 155
 Db 339 TCATTGGAA-----AATTATAAGATATGATCATATAAT--- 371
 Qy 156 IleAspPheTyrGlnAspThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysAsnLeu 175
 Db 372 -----AATAATAATAATAGTAATAAT 392
 Qy 176 LeuAsnLysCysThrGlyPheLysAsnGlnLysPheLysPheLeuAsnHisSerMetAla 195
 Db 393 ATGGATGTGATTCATTTATTAATAGTAGAAGAAATAATCATACCTTCATTGAT 452
 Qy 196 CysAspLeuGlnAlaLeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeu 215
 Db 453 ---GATAAATATTCATTACAATATTTAAA----- 479
 Qy 216 AsnSerAspGluTyrPheLysLeuAlaProAspIleGlySerAspPheValLeuSerLys 235
 Db 480 AATTTTATGATCATGATAAATTA-----ATTTTTTAAACAAT 518
 Qy 236 LysProLysLysTyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArg 254
 Db 519 ATGATTTCTAAATATATA-----TATGCAATTAGTGTATAAACATTTTATGAACGG 569
 RESULT 5
 CK573341
 LOCUS 1381 Plasmodium yoelii liver stage LCM cDNA library Plasmodium
 DEFINITION yoelii cDNA, mRNA sequence.
 ACCESSION CK573341
 VERSION CK573341.1 GI:44827431
 KEYWORDS EST.
 SOURCE Plasmodium yoelii
 ORGANISM Plasmodium yoelii
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 REFERENCE 1 (bases 1 to 894)
 AUTHORS Sacci,J.B. Jr., Ribeiro,J.M.C., Huang,F., Alam,U., Russell,J.A.,
 Blair,P.L., Witney,A., Carucci,D.J., Azad,A.F. and Aguiar,J.C.

Qy	175	LeuLeuAenLysCysThrGlyPhelyAenGlnLysPhelyShelleAenHisSerMet	194
Db	447	ATAAGTAACAAAATAACATTTCTGAGAATTATAGATTTTATTATTGCTAAT--AAATTTA	503
Qy	195	AlaCysAspLeuGlnAlaLeuAapTyrLeuMetLysArgTyrAspValaAenIleTyrSer	214
Db	504	AAATGT-----GATCATATTGTTATTAATAAACCATTAACCTTTATTATTATG	548
Qy	215	LeuAenSer-----AspGlu-----TyrPhelyLeu	223
Db	549	ATTAAATTCATATTATTATTTTAAATATAAAATATGACAAAATAACATTTTATAAAATC	608
Qy	224	AlaProAspIleGlySerAspPheValLeuSerLysLysPro-----LysLysTyrIle	241
Db	609	GTG-----TCTATAATTTTAAATCGNAAAACCCAGAAAATCGAAATATATATT	653

RESULT 6

BM167106

LOCUS

DEFINITION

EST569629 PyBS Plasmodium yoelii yoelii cDNA clone pYCOC76 5' end, mRNA sequence.

BM167106

BM167106.1 GI:17300338

EST.

Plasmodium yoelii yoelii

Plasmodium yoelii yoelii

Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

1 (bases 1 to 774)

Carlton,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Vaidya,A.B., Fraser,C.M. and Carucci,D.J.

Plasmodium yoelii EST project at TIGR

Unpublished (2001)

Contact: Jane Carlton

Parasite Genomics Group

The Institute for Genomic Research

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-530-9319

Fax: 301-838-0208

Email: carlton@tigr.org

For clone info, please contact the Malaria Research and Reference Reagent Resource Center, ATCC

<http://www.malaria.mr4.org/mr4pages/index.html>

Seq primer: ADF.

Location/Qualifiers

1..774

/organism="Plasmodium yoelii yoelii"

/mol_type="mRNA"

/strain="17XL"

/sub_species="yoelii"

/db_xref="taxon:73239"

/clone="PYCOC76"

/dev_stage="Asexual blood stages"

/lab_host="E. coli XL-1 Blue"

/clone_lib="PyBS"

/note="Vector: At 20-25% parasitemia, blood was collected from BALB/cByJ mice infected with Py17XL parasites, and leukocytes removed by passage over microcrystalline cellulose columns. Total RNA was isolated using the guanidinium isothiocyanate method, and mRNA isolated using oligo(dT)-cellulose chromatography. First strand cDNA synthesis was completed using a 50-base primer and reverse transcriptase in the presence of 5-methyl dCTP. After second strand synthesis, uneven termini were treated with Pfu DNA polymerase and EcoRI adaptors ligated to the blunt ends. The sample was cleaved with XhoI and separated on a Sephacryl S-500 column. Size-fractionated cDNA was precipitated and ligated to HybriZAP arms directionally using EcoRI-XhoI cleaved arms. After packaging, the phagemid vector (pAD-GAL4) was excised from the HybriZAP vector and plasmid DNA isolated."

FEATURES

source

ORIGIN

Alignment Scores:

```

Pred. No.:      0.0150      Length:      774
Score:          111.50      Matches:      57
Percent Similarity: 36.10%      Conservative: 43
Best Local Similarity: 20.58%      Mismatches: 80
Query Match:      7.10%      Indels:      97
DB:              4          Gaps:      11

US-10-735-419-10 (1-294) x BM167106 (1-774)

Qy 15 SerLeuLysAsnIleAspTyrLysArgLeuProLysGlnPheAspValPheArgCysAsn 34
Db 108 AACATAGGCAAGTTAGCATCAAGAACTAGCAAAACAA---AAGGTAACATGAGTAAT 164
Qy 35 GlnPheTyrPheGluAspArgTyrPheValGlyLysAspValLysTyrValPheAsn 54
Db 165 -----ATATACAAATTTGTACGACATAGGGCTTACANTGATACATAATTTTAAC 218
Qy 55 ProPheValPhePheGluGlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGlu--- 73
Db 219 ACATATATA-----GAAACAACATGCAGAGATGAAGACAA 254
Qy 74 -----TyrAsnIleGluAsnIleValCysSerThrIleAsn----- 85
Db 255 TTACATTTAGAAATAATTCACACAAAAATAATATATTTAGTGATAGAAATTAATATG 314
Qy 86 -----LeuGluTyr-----IleAsp 90
Db 315 GTTGATTATTAATAAAAAAACCATTAATATACATAATGTATGCAAAAGTATTATGAT 374
Qy 91 GlyPheGlnPheValAspAsnPheGluTyrPheSerAspAlaPheLeuGlyHisGlu 110
Db 375 TCGTATAAAAAAAGATAT-----GTATTTTAAAAAARACA 413
Qy 111 IleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIleTyrAsnArg 130
Db 414 ATATTAGAATGTTTGAAGCATATATTTTATTTGTTTGCATATGAA----- 461
Qy 131 GlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGlyTyrLysSer 150
Db 461 ----- 461
Qy 151 IleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAlaPheAspAsn 170
Db 462 -----AATTTATTAAACATTTGAAGAA 482
Qy 171 AsnLysLysAsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPheLysPheIle 190
Db 483 AATGATCAGATGTATTAATAATATATATTAATATATATTTTAAACAGATTTAAATTTAAG 542
Qy 191 AsnHisSerMetAla-----CysAspLeuGlnAlaLeu 201
Db 543 GATTATAATGATGATCTCTTTTACTGTTTAAATCTTTAAATTTATATATATATGCAVTT 602
Qy 202 AspTyrLeuMetLysArgTyrAspValAsn-----IleTyr 213
Db 603 GAAATTTTACAAAAAATTTATATACATAATATTTTATTTAAAAAATGCTGTATAT 662
Qy 214 SerLeuAsnSerAspGluTyrPheLysLeuAlaProAspIleGlySerAspPheValLeu 233
Db 663 AGTATGATATCGAGTTGTTGTTTTTATATACGCAACAAATATA---AATGAATTTTGTTTA 719
Qy 234 SerLysLysProLysTyrIleAsnAspIleLeuIleProAspLysTyr 250
Db 720 AAAAAAAGGGAATAGATATACCAATTTTAAATGTATTATTAGAGCCTAC 770

```

RESULT 7

```

AZ935754      1123 bp      DNA      linear      GSS 21-DEC-2001
LOCUS      AZ935754
DEFINITION      ODG282 Oikopleura dioica Shotgun Library Oikopleura dioica genomic
                  clone G282, genomic survey sequence.
ACCESSION      AZ935754
VERSION      AZ935754.1 GI:17975925

```

KEYWORDS

```

SOURCE      Oikopleura dioica
ORGANISM      Oikopleura dioica
REFERENCE      Eukaryota; Metazoa; Chordata; Urochordata; Appendicularia;
AUTHORS      Oikopleuridae; Oikopleura.
              1 (bases 1 to 1123)
              Seo,H.C., Kube,M., Edvardsen,R.B., Jensen,M.F., Beck,A., Spriet,E.,
              Gorsky,G., Thompson,E., Lehrach,H., Reinhardt,R. and Chourrout,D.
              Miniature genome in the marine chordate Oikopleura dioica
              Science 294 (5551), 2506 (2001)
TITLE      Oikopleura dioica
JOURNAL      21625354
MEDLINE      11752568
PUBMED
COMMENT      Contact: Seo HC
              Sars International Centre for Molecular Marine Biology
              Bergen High Technology Centre, Thormohlenstg. 55, N-5008 Bergen,
              Norway
              Tel: 55 58 4389
              Fax: 55 58 4305
              Email: hee-chan.seo@sars.uib.no
              Insert Length: 1123 Std Error: 0.00
              Seq primer: T7 and V4 (ATCCAACGGCTTGGGAGCTCT)
              Class: shotgun.

```

FEATURES

```

source
1..1123
/organism="Oikopleura dioica"
/mol_type="genomic DNA"
/db_xref="taxon:34765"
/clone="G282"
/sex="male"
/clone_lib="Oikopleura dioica Shotgun Library"
/notes="Organ: sperm; Vector: pGEM-T Easy; The shotgun
library was prepared using the EcoRV digested pGEM-T Easy
vector (Promega) by Dr Hee-Chan Seo."

```

ORIGIN

```

Alignment Scores:
Pred. No.:      0.0849      Length:      1123
Score:          107.00      Matches:      68
Percent Similarity: 34.38%      Conservative: 42
Best Local Similarity: 21.25%      Mismatches: 98
Query Match:      6.81%      Indels:      112
DB:              8          Gaps:      15

US-10-735-419-10 (1-294) x AZ935754 (1-1123)

Qy 22 LysArgLeuProLysGlnPheAsp---ValPheArgCysAsnGlnPheTyrPheGluAsp 40
Db 91 AAGCTCATCCCTAAAAATTTGTGACATTATATACAGTCTCCAGAAATTCATTTTAATGAA 150
Qy 41 ArgTyrPheValGlyLysAspValLysTyrValPhePheAsnProPheValPheGlu 60
Db 151 GAA-----TTGAAAAAAGATGGAAACCAATGCCACTTCCACCACCGAAATTTTCAA 204
Qy 61 -----GlnTyrTyrThrSerLysLysLeuIleGlnAsnGluGlu 73
Db 205 AATTGTGATGACTCAAGGTCTTCAATATTTTACAGCA---TTTGGCAGCGCAAAACAATTT 261
Qy 74 TyrAsn-----IleGluAsnIle 79
Db 262 TATCGTCAATTTATCTGCTGTGCGATGGATGCCAGTACGTCCCGGTTCTAGCTCGATC 321
Qy 80 ValCysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnPheGlu 99
Db 322 CTTTGTCTCTTTTGAAGAAGATTCATCAT----- 354
Qy 100 LeuTyrPheSerAspAlaPheLeuGlyHisGluIleIleLysLysLeuLysAspPhe 119
Db 355 -----CATGAG----- 360
Qy 120 AlaTyrIleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMet 139
Db 360 ----- 360

```

```

QY 140 CysAlaThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyr 159
Db 361 -----GTACGGTAGGGCATGTGCTGCTATATTCCTCGGTATTCGAT----- 402

QY 160 GlnAspThrAsnAsnLeuTyrAlaPheAsp-----AsnAsnLysLysAsnLeuLeuAsn 177
Db 403 ---GATATATTTACTATCTGCGGAAGACAAATTTGACCATCATATAATTTTAAAC 459

QY 178 LysCysThrGlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAsp 197
Db 460 TCTTGG-----GACCCGGATATAAAATTTACAAAAGAATCGCGAGCTTCAGGT 507

QY 198 ---LeuGlnAlaLeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsn 216
Db 508 CTTTGAATTTTTCGATCGCATCGTTGAA---TAGATAAAACCACTAAACCTTGA 564

QY 217 SerAspGluTyrPheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLys 236
Db 565 ATGGCAACATTTTCATAAAATACAGTAAATAGATCACTCCAGAAATTTTCATTCGATTACC 624

QY 237 ProLysLysTyrIleAsnAspIleLeuLeuProAspLysTyrAlaGlnGluArgTyrTyr 256
Db 625 CGGAAGACCAAAATATGATTTATGTTGGTGAGATTCATCGAATTAATAACACGTGC 684

QY 257 GlyLysLysSerArgLysGluAsnLeuHis----- 267
Db 685 ACCAAAGTACGCACCTAGAAAATCTCTTCATGATCTCAGCAACAATTTATTCGAAC 744

QY 268 -----TyrLysLeuIleLysAsp-----Leu 274
Db 745 GCCTATCTCCACACTAGCTTTTCACAAATATAAGCAAGTTAAGATAGTGGCTTTAA 804

QY 275 IleArgLeuProSer-----AspIleLysHisTyrLeuLysGluLysTyrAlaAsn 291
Db 805 CTCCTGTAAGAAGATCCCATATAGACATAGACATATCTCGAACTTCGACTTTCACAAAT 864

RESULT 8
CA855294 666 bp mRNA linear EST 17-DEC-2002
LOCUS PfESToc3sf10.y1 Plasmodium falciparum 3D7 gametocyte cDNA library
DEFINITION Plasmodium falciparum 3D7 cDNA 5', mRNA sequence.
ACCESSION CA855294
VERSION CA855294.1 GI:27157046
KEYWORDS EST.
SOURCE Plasmodium falciparum 3D7
ORGANISM Plasmodium falciparum 3D7
REFERENCE 1 (bases 1 to 666)
AUTHORS Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
Marra,M., Hillier,L., Martin,J., Wylie,T., Dante,M., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Jentes,E., Ronko,I.,
Tsagarisshvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,
Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
Waterston,R., Wilson,R. and Sibley,D.
WashU Plasmodium EST Project
Unpublished (2001)
TITLE Contact: L. David Sibley
JOURNAL WashU Plasmodium EST Project
COMMENT Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Library was constructed by R. Haywood. DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: L. David Sibley
(sibley@borcim.wustl.edu), Washington University
Seq primer: -400P from Gibco
High quality sequence stop: 453.
Location/Qualifiers
FEATURES
source
1..666
/organism="Plasmodium falciparum 3D7"
/mol_type="mRNA"

```

```

/db_xref="taxon:36329"
/dev_stages="gametocyte (stage III-V)"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="Plasmodium falciparum 3D7 gametocyte cDNA
library"
/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
XhoI; The library was constructed by R Haywood. cDNAs were
synthesized from gametocyte poly(A)+ RNA by oligo d(T)
priming, size-selected and directionally cloned into the
EcoRI (5' end) to XhoI (3' end) sites of the Uni-ZAP XR
lambda vector (Stratagene). The primary library was mass
excised as phagemid using the ExAssist helper phage
(Stratagene). Clones were mass excised using the ExAssist
helper phage (Stratagene), the phagemids were precipitated
with PEG 8000 and extracted with phenol/chloroform.
Phagemid DNA was electroporated into DH10B cells. Clone
Availability: David Sibley, Washington University."

ORIGIN
Alignment Scores:
Pred. No.: 0.0491 Length: 666
Score: 106.50 Matches: 52
Percent Similarity: 39.90% Conservative: 29
Best Local Similarity: 25.62% Mismatches: 85
Query Match: 6.78% Indels: 37
DB: 6 Gaps: 9

US-10-735-419-10 (1-294) x CA855294 (1-666)
QY 110 GluIleLysLysLysLysAspPhePheAlaTyrIleLysTyr-----AsnGlu 126
Db 9 GAATGTATAAATTTACTTTAAATAATTCATTACTAATATTATACATACATACATACAC 68

QY 127 IleTyrAsnArgGlnArgIleThrSerGlyValTyr-----MetCys 140
Db 69 CTATTAAATACAGATCATATACAAAGATAAATAAATAAACAAGAAAGTACAAAGT 128

QY 141 AlaThrAlaValAlaLeuGlyTyrLys-----SerIleTyrIleSerGly 155
Db 129 GCTCATGATGATTTTACGTCCTTATAGAAATAAATAAGGATGATAATATTTTGTATCATCAT 188

QY 156 IleAspPheTyrGlnAspThrAsnAsnLeuTyrAla-----PheAspAsnAsnLysLys 173
Db 189 TTTAAACATTTATGAGATGAGATGAGATTTATATATTTATTTTAAATAAATAATTTATAA 248

QY 174 AsnLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPheLysPheIleAsnHisSer 193
Db 249 AATAAA-----AAGGATCAACAAATAATTTTAGTAAATATGAGA 287

QY 194 MetAlaCysAspLeu-----GlnAlaLeuAspTyrLeuMetLysArg--- 207
Db 288 AAAAAATAAGATTTTAAATGATTTATATCATGATTTTAAATAATATATATTTAAACGAAAT 347

QY 208 -----TyrAspValAsnIleTyrSerLeuAsnSerAspGluTyrPheLysLeuAla 224
Db 348 TCATCTTTATACATATATGATAATATTAATATATTTGATGAT---GTTTTTAATTTACTT 404

QY 225 ProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAspIle 244
Db 405 TCTTTAATGCTGCTATATATATATATATATATATATATATATATATATATATATATATAT 464

QY 245 LeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSerArgLeuLysGlu 264
Db 465 TCTTTTATTTGGGAATATATATATATATATATATATATATATATATATATATATATATAT 524

QY 265 AsnLeuHisTyrLysLeuIleLysAspLeuIleArgLeuProSerAspIleLysHisTyr 284
Db 525 AATTTATATCTTAAATATGATACAGAAATTTATTTA-----GATATGAAGAACAAT 572

285 LeuLysGlu 287
:::|
573 ATTAAGAA 581

```

```
RESULT 9
BJ371051          678 bp  mRNA  linear  EST 08-MAR-2002
LOCUS             BJ371051 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION        BJ371051 Dictyostelium discoideum cDNA clone ddc56p08 5', mRNA sequence.
ACCESSION         BJ371051
VERSION           BJ371051.1 GI:19280434
KEYWORDS
SOURCE            Dictyostelium discoideum
ORGANISM          Dictyostelium discoideum
REFERENCE         1 (bases 1 to 678)
AUTHORS           Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE             Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL           Unpublished (2002)
COMMENT           Contact: Tadasu Shin-i
                   Center For Genetic Resource Information
                   National Institute of Genetics
                   1111 Yata, Mishima, Shizuoka 411-8540, Japan
                   Tel: 81-559-81-6856
                   Fax: 81-559-81-6855
                   Email: tshini@genes.nig.ac.jp.
                   Location/Qualifiers
FEATURES           source
                   1..678
                   /organism="Dictyostelium discoideum"
                   /mol_type="mRNA"
                   /strain="AX4"
                   /db_xref="taxon:44689"
                   /clone="ddc56p08"
                   /sex="mat A"
                   /dev_stage="Culmination stage"
                   /clone_lib="Dictyostelium discoideum cDNA library, CF"

ORIGIN
Alignment Scores:
Pred. No.:      0.0502      Length:      678
Score:          106.50      Matches:    59
Percent Similarity: 35.40%  Conservative: 38
Best Local Similarity: 21.53% Mismatches: 80
Query Match:    6.78%      Indels:    97
DB:             4          Gaps:      14

US-10-735-419-10 (1-294) x BJ371051 (1-678)

Qy 46 LysAspValLysTyrValPhePheAsnPropheValPheGluGlnTyrTyrThrSer 65
Db 58 AAAGATATAAAATTAATA-----ATTGAACAATTAAGGGAATTA 96
Qy 66 LysLysLeuIle-----GlnAsnGluGluTyrAsn----- 75
Db 97 AAACAATCATTTTAAATATAATATAGCGAAATATAGATATGATATCAAGTAACA 156
Qy 76 -----lGluAsnIleVal-----Cys 81
Db 157 ATAACATCAATTTTAAATTAATTAATGAAATGTTTGTAGATAATGGTATTATGCAATTC 216
Qy 82 SerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheValAspAsnGluLeuTyr 101
Db 217 ACAGAAATTGATAACAATTAATTTGAAAGTTTTCTTTTACCACAAATTTA----- 267
Qy 102 PheSerAspAlaPheGlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyr 121
Db 268 TTATCAGAAATCATTTATA-----ACATTAACAACATTTATTCAGATAATTAATTA 318
Qy 122 IleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAla 141
Db 319 AATAATAGTAATAGTAATAATAATAATAATAGTAAA----- 351
Qy 142 ThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAsp 161
Db 352 -----AATAATGAGAAATTTATATTCACAAAAATTAACCTTTT----- 390
```

```
Qy 162 ThrAsnAsnLeuTyrAlaPhe-----AspAsnAsnLysLysAsnLeuLeuAsnLysCys 179
Db 391 TTTAATATTTTATATAAAATTTATAAATGATAATAATAGTAATAATTTTATCAAGTTAATA 450
Qy 180 ThrGlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGln 199
Db 451 ATTGAATTTCAAAATAGAAAATTTTAAA-----AAACATATAATAATAATTTGATTAGAA 504
Qy 200 AlaLeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGlu 219
Db 505 ATTCAATGAACAATTT-----TGGAAATTTATAATAAATCAATTT----- 543
Qy 220 TyrPheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLys 239
Db 543 ----- 543
Qy 240 TyrIleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLys 259
Db 544 -----TGTTTTGATACCAGATAAAATGTAGATCAATTTCAATATTTATATAAAAT 591
Qy 260 SerArgLeuLysGluAsnLeuHisTyrLysLeuIleLysAspLeuIleArgLeuProSer 279
Db 592 AATAGAATTAGG---AATATTCAAATCACCATT-----TCA 624
Qy 280 AspIleLysHisTyrLeuLysGluLysTyrAlaAsnLysAsn 293
Db 625 TCAGTTAAGATTATATTTTAGAATTTCTTTACCAATAATAAT 666

RESULT 10
BJ365708          833 bp  mRNA  linear  EST 08-MAR-2002
LOCUS             BJ365708 Dictyostelium discoideum cDNA library, CF Dictyostelium
DEFINITION        BJ365708 Dictyostelium discoideum cDNA clone ddc36g07 5', mRNA sequence.
ACCESSION         BJ365708
VERSION           BJ365708.1 GI:19275010
KEYWORDS          EST.
SOURCE            Dictyostelium discoideum
ORGANISM          Dictyostelium discoideum
REFERENCE         1 (bases 1 to 833)
AUTHORS           Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
TITLE             Full length cDNA of Dictyostelium discoideum at the culmination
stage
JOURNAL           Unpublished (2002)
COMMENT           Contact: Tadasu Shin-i
                   Center For Genetic Resource Information
                   National Institute of Genetics
                   1111 Yata, Mishima, Shizuoka 411-8540, Japan
                   Tel: 81-559-81-6856
                   Fax: 81-559-81-6855
                   Email: tshini@genes.nig.ac.jp.
                   Location/Qualifiers
FEATURES           source
                   1..833
                   /organism="Dictyostelium discoideum"
                   /mol_type="mRNA"
                   /strain="AX4"
                   /db_xref="taxon:44689"
                   /clone="ddc36g07"
                   /sex="mat A"
                   /dev_stage="Culmination stage"
                   /clone_lib="Dictyostelium discoideum cDNA library, CF"

ORIGIN
Alignment Scores:
Pred. No.:      0.114      Length:      833
Score:          104.50      Matches:    62
Percent Similarity: 38.66%  Conservative: 42
Best Local Similarity: 23.05% Mismatches: 88
Query Match:    6.65%      Indels:    77
DB:             4          Gaps:      14

US-10-735-419-10 (1-294) x BJ365708 (1-833)
```


Qy 71 AenGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGluTyrIleAsp 90
 Db 14 AATAAATTTTATTCATACCATTAATTAATTAATTAATTAATTAATTCGTGT----- 61
 Qy 91 GlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeuGlyHisGlu 110
 Db 62 GGTGCTGTTTTTATCCAGGAGAGGAGCATATTTTGAATATAGGATTCATAATGAAAAA 121
 Qy 111 IleIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIleTyrAsnArg 130
 Db 122 TTTATTTTCAACACTAGTGATGATTT-----AAAATTCACACATGCA 163
 Qy 131 GlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAla-----LeuGly 147
 Db 164 AGAGAGATTAATTCGTGATGAATAAGTTTACAAACGTTGTAGGTAAATTAATTAAC 223
 Qy 148 TyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAla 167
 Db 224 AAACCAATAGATATAATTAATTTGGGATTTTCATCTTGACCCAGACTCAGTTACTTTT 283
 Qy 168 PheAspAsnAsnLysLysLeuLeuAsnLysCysThrGlyPheLysAsnGlnLysPhe 187
 Db 284 TTTGAATCTGAA-----ACAGAGTTTGTGTAACCAAAAAATTC 319
 Qy 188 LysPheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArg 207
 Db 320 CAGTATGTAAT-----GGATATTTGATAAATCTT 349
 Qy 208 TyrAspVal-----AsnIleTyr-----SerLeuAsnSerAspGluTyr 220
 Db 350 TGTGAAGTATAATTTGTTATTTTGGCCATGGATGGTTTCACITGTTAGGATTAATAT 409
 Qy 221 PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr 240
 Db 410 TTT-----GGAGATATTAACCAAGATCTATTACTTCGTAATAATGAAAAATCAACAA 460
 Qy 241 IleAsnAspIle-----Leu 245
 Db 461 GTAAACAGATATATGAAGGTATTAATAACAGATTAATTCGCTGTAATAATAAAT 520
 Qy 246 IleProAspLysTyrAlaGlnGluArgTyrGlyLysLys-----SerArgLeuLys 263
 Db 521 CATCCAGAAAAAATGTCATGAGAGGATATTAACCAACAAAGTTTTTTCCGTATCAAA 580
 Qy 264 GluAsnLeuHisTyrLysLeuIle-----LysAspLeuIleArg--- 276
 Db 581 ---AATATTCATGAAAAACCAAGTTTATGAGGTGATATCAACCAAGATTTTATTACGTCGT 637
 Qy 277 -----LeuProSerAspIleLysHis---TyrLeu 285
 Db 638 AAAAATAAATAATCCAGAGAAACAAGTATATGAAAGTGATATTAACCAAGATTTTATTA 697
 Qy 286 LysGluLysTyrAlaAsnLysAsnArg 294
 Db 698 CGTCGTAATATTAATGATAAATCTCAA 724

RESULT 11
 AZ679344/c
 LOCUS
 DEFINITION ENTMB86TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
 genomic, genomic survey sequence.
 ACCESSION AZ679344
 VERSION AZ679344.1 GI:11816490
 KEYWORDS GSS.
 SOURCE Entamoeba histolytica
 ORGANISM Entamoeba histolytica
 Eukaryota; Entamoebidae; Entamoeba.
 REFERENCE 1 (bases 1 to 873)
 AUTHORS Loftus, B., Van Aken, S. and Fraser, C.
 TITLE Determination of clone end sequences from Entamoeba histolytica
 HMI:IMSS sheared DNA library
 JOURNAL Unpublished (2000)
 COMMENT Contact: Brendan J Loftus

Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0208
 Fax: 301 838 3543
 Email: bjoftus@tigr.org

Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
 DNA library
 Seq primer: M13-Reverse
 Class: Shotgun
 High quality sequence start: 39
 High quality sequence stop: 856.

FEATURES source

1..873

/organism="Entamoeba histolytica"
 /mol_type="genomic DNA"
 /strain="HMI:IMSS"
 /db_xref="taxon:5759"
 /clone_lib="Entamoeba histolytica Sheared DNA"
 /note="Vector: pHOS1; Site 1: Bst I; Constructed at The
 Institute for Genomic Research (TIGR), Rockville, MD.
 Genomic DNA isolated from broth cultures of E. histolytica
 using a method described by Clark and Diamond (Clark,
 C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
 method for isolate identification. Exp. Parasitol.
 77:450.). The DNA was mechanically sheared to give a
 tight size distribution (~2 kb). The v + i method used for
 the library construction is described in detail in Smith,
 H.O. and Venter, J.C. (Making small insert libraries for
 whole genome shotgun sequencing projects. In Genome
 Sequencing: A Practical Approach, eds. M. Vaudin and B.
 Borell, Oxford University Press, 1999)."

ORIGIN

Alignment Scores:
 Pred. No.: 0.121 Length: 873
 Score: 104.50 Matches: 62
 Percent Similarity: 40.46% Conservative: 44
 Best Local Similarity: 23.66% Mismatches: 80
 Query Match: 6.65% Indels: 76
 DB: 8 Gaps: 15

US-10-735-419-10 (1-294) x AZ679344 (1-873)

Qy 16 LeuLysAsnIleAspTyrLysArgLeuProLysGlnPheAspValPhe----- 31
 Db 813 ATTAATCAATAAATTACAAAAAGTGTAACAAA---TTTAATATATTCAAAAAGTGAAG 757
 Qy 32 ArgCysAsnGln-----PheTyrPheGluAspArgTyrPheValGlyLysAspValLys 49
 Db 756 AGGTGCCAAAATATCTTTTATTACTTCAAAAATATTTAT-----AATAAGAAA 709
 Qy 50 TyrValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLysLeuIle 69
 Db 708 AATTATCAAAATAACGAGATATAATACTTCAAAAATACGTATGAAATGAAATATACG 649
 Qy 70 GlnAsnGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeuGluTyrIle 89
 Db 648 AAAAGAGATGAAAAATAAC---GAAGATATAATACTTCAAAATATC----- 607
 Qy 90 AspGlyPheGlnPheValAspAsnPheGluLeu-----TyrPheSerAsp 104
 Db 606 -----TTCATTATGAAATTTTCGAAAAATAACGAGATATAATACTTTC----- 565
 Qy 105 AlaPheLeuGlyHisGluIleIleLysLysLeuLysAspPhePheAlaTyrIle---Lys 123
 Db 564 -----AAAATAAGAAATATCTTAAGATAGTTCAAAAA 532
 Qy 124 TyrAsnGluIleTyrAsnArgGlnArgIleThrSer-GlyValTyrMetCysAlaThrAl 143
 Db 531 TATCCAAACATCGTTCATTTAATTAATGGAATACCTCAAAATATCTGTATCTGTTCTAAAA 472
 Qy 143 aValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAs 163

```

Db 471 TATCTGT-----ATCTATTCTAAATAATAATAATAATCTATTATTATAGA 427
Qy 163 nAsnLeuTyralaPheAsnLysLys-----AsnLeuLe 176
Db 426 AAAATGGTATTTTAAACCAATAATAAAAAATGTTATATAAAACAATAGTAA 367
Qy 176 uAsnLysCys---ThrGlyPheLysAsnGlnLysPheLysPheLeAsnHisSerMetAl 195
Db 366 TAAAAATGTTATTATTATCCCAAAGGACCAATTTCCACTGTGTATATATGATGATC 307
Qy 195 a-----CysAspLeuGlnAlaLeuAspTyLeuMetLysArgTyRAspValAsnIleTy 213
Db 306 ACATATTGTCAATTTGACCATAGAAAAATATTATAAAGATATAATTAATTCAAT 247
Qy 213 rSerLeuAsnSerAspGluTyRPhelLysLeuAlaProAspIleGlySerAspPheValLe 233
Db 246 AAATTTAAATTTAAAT-----TT 229
Qy 233 uSerLysLysProLys-----LysTyRileAsnAspIleLe 245
Db 228 AAATAAATGAATAATAATAATTTAAATAGAAAAATAATAAAGTAATTAATTAAAGTATT 169
Qy 245 uile 246
Db 168 AATA 165

RESULT 12
AZ546559/c
LOCUS
DEFINITION ENTDR06TF Entamoeba histolytica Sheared DNA linear GSS 14-NOV-2000
genomic, genomic survey sequence.
ACCESSION AZ546559
VERSION AZ546559.1 GI:11169417
KEYWORDS GSS.
SOURCE Entamoeba histolytica
ORGANISM Entamoeba histolytica
Eukaryota; Entamoebidae; Entamoeba.
REFERENCE 1 (bases 1 to 924)
Loftus,B., Van Aken,S. and Fraser,C.
Determination of clone end sequences from Entamoeba histolytica
HMI:IMSS sheared DNA library
Unpublished (2000)
JOURNAL
COMMENT Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0208
Fax: 301 838 3543
Email: b.loftus@tigr.org
Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
DNA library
Seq primer: M13-Forward
Class: shotgun
High quality sequence start: 94
High quality sequence stop: 921.
FEATURES
source
1..924
/organism="Entamoeba histolytica"
/mol_type="genomic DNA"
/strain="HMI:IMSS"
/db_xref="taxon:5759"
/clone_lib="Entamoeba histolytica Sheared DNA"
/notes="Vector: pHOSt1; Site_1: Bst I; Constructed at The
Institute for Genomic Research (TIGR), Rockville, MD.
Genomic DNA isolated from broth cultures of E. histolytica
using a method described by Clark and Diamond (Clark,
C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
method for isolate identification. Exp. Parasitol.
77:450.). The DNA was mechanically sheared to give a
tight size distribution (~2 kb). The v + i method used for
the library construction is described in detail in Smith,
H.O. and Venter, J.C. (Making small insert libraries for

```

whole genome shotgun sequencing projects. In Genome Sequencing: A Practical Approach, eds. M. Vaudin and B. Barrell, Oxford University Press, 1999).

ORIGIN

Alignment Scores: 0.131 Length: 924
Pred. No.: 104.50 Matches: 62
Score: 104.50 Matches: 62
Percent Similarity: 40.46% Conservative: 44
Best Local Similarity: 23.66% Mismatches: 80
Query Match: 6.65% Indels: 76
DB: 8 Gaps: 15

US-10-735-419-10 (1-294) x AZ546559 (1-924)

Qy 16 LeuLysAsnIleAspTyRLeuArgLeuProLysGlnPheAspValPhe----- 31
Db 738 ATTAAATCAATAATAATACAAAAAGTGTAACAAA---TTTAAATATATCAAAAAAGTGAAG 682

Qy 32 ArgCysAsnGln-----PheTyRPhelGluAspArgTyRPhelValGlyLysAspValLys 49
Db 681 AGTGCCAAAATATCTTTTATTACTTCAAAAATATTAT-----AATAAGAAA 634

Qy 50 TyrValPhePheAsnPropheValPhePheGluGlnTyRtyrThrSerLysLysLeuLe 69
Db 633 AATTATGAAAAATAACGAAGATATAATACTTCAAAAATACGTATGAAATGAAAAATATACG 574

Qy 70 GlnAsnGluGluTyRAsnIleGluAsnIleValCysSerThrIleAsnLeuGluTyRile 89
Db 573 AAAAGAGATGAAAAATAAC---GAAGATATAATACTTCAAAAATATC----- 532

Qy 90 AspGlyPheGlnPheValAspAsnPheGluLeu-----TyrPheSerAsp 104
Db 531 -----TTCATTATGAATTTTCGAAATAACGAAGATATAATACTTCTTC----- 490

Qy 105 AlaPheLeuGlyHisGluIleLeLysLysLeuLysAspPheAlaTyRile---Lys 123
Db 489 -----AAAAAAGAAAAATATCTAAGAAATAGTTCAAAAA 457

Qy 124 TyrAsnGluIleTyRAsnArgGlnArgIleThrSer-GlyValTyRMetCysAlaThrAl 143
Db 456 TATCCAAACAATCGTTCATTTTAAATGGAATACCTCAAAAATATCTGTATCTGTCTCTAAAAA 397

Qy 143 aValAlaLeuGlyTyRlyssSerIleTyRileSerGlyIleAspPheTyRGlNAspThrAs 163
Db 396 TATCTGT-----ATCTATTCTAAAAATATAAATATAATCTATTATTATAGA 352

Qy 163 nAsnLeuTyRAlaPheAspAsnLysLys-----AsnLeuLe 176
Db 351 AAAATGCTATTTTAAAAACAATAATAAAAAATAAAAAATGTTATATATAAAACAATAGTAA 292

Qy 176 uAsnLysCys---ThrGlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAl 195
Db 291 TAAAAAATGTTATTATTATCCCAAAAGGACAGATTTTCCCACTGTGTATATATGATGATC 232

Qy 195 a-----CysAspLeuGlnAlaLeuAspTyRLeuMetLysArgTyRAspValAsnIleTy 213
Db 231 ACATATTGTCAATTTGACCATAGAAAAATATTATAAAGATATAATTAATTCAAT 172

Qy 213 rSerLeuAsnSerAspGluTyRPhelLysLeuAlaProAspIleGlySerAspPheValLe 233
Db 171 AAATTTAAATTTAAAT-----LysTyRileAsnAspIleLe 154

Qy 233 uSerLysLysProLys-----LysTyRileAsnAspIleLe 245
Db 153 AAATAAATGAATAAAAAATATTTTAAATAGAAAAATAAATAAAGTAATTAATTAAAGTATT 94

Qy 245 uile 246
Db 93 AATA 90

RESULT 13
AZ684114/c

```

LOCUS       AZ684114                      973 bp      DNA      linear      GSS 14-DEC-2000
DEFINITION  ENTUQ74TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, genomic survey sequence.
ACCESSION   AZ684114
VERSION     AZ684114.1  GI:11821260
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
  1 (bases 1 to 973)
  Loftus,B., Van Aken,S. and Fraser,C.
  Determination of clone end sequences from Entamoeba histolytica
  HMI:IMSS sheared DNA library
  Unpublished (2000)
  Contact: Brendan J Loftus
  Department of Eukaryotic Genomics
  The Institute for Genomic Research
  9712 Medical Center Dr., Rockville, MD 20850, USA
  Tel: 301 838 0208
  Fax: 301 838 3543
  Email: bjloftus@igr.org
  Clones are derived from the Entamoeba histolytica HMI:IMSS sheared
  DNA library
  Seq primer: M13-Forward
  Class: shotgun
  High quality sequence start: 101
  High quality sequence stop: 281.
  Location/Qualifiers
    1..973
      /organism="Entamoeba histolytica"
      /mol_type="genomic DNA"
      /strain="HMI:IMSS"
      /db_xref="taxon:5759"
      /clone_lib="Entamoeba histolytica Sheared DNA"
      /note="Vector: pROS1. Site 1: Bst I; Constructed at The
            Institute for Genomic Research (TIGR), Rockville, MD.
            Genomic DNA isolated from broth cultures of E. histolytica
            using a method described by Clark and Diamond (Clark,
            C.G., and Diamond, L.S. (1993) Entamoeba histolytica: a
            method for isolate identification. Exp. Parasitol.
            77:1450.). The DNA was mechanically sheared to give a
            tight size distribution (~2 kb). The v + i method used for
            the library construction is described in detail in Smith,
            H.O. and Venter, J.C. (Making small insert libraries for
            whole genome shotgun sequencing projects. In Genome
            Sequencing: A Practical Approach, eds. M. Vaudin and B.
            Borell, Oxford University Press, 1999)."
```

Db	545	GCAGATATAATCCCTTCAAAAATACGTTATGAAATGAAACAATATATACGAAAGAGATGAAAT	486
Qy	76	IleGluAsnIleValCysSerThrIleAsnLeuGluTyrIleAspGlyPheGlnPheVal	95
Db	485	ACCGAAGATATAATCTCTCAAAATATA-----TTCATTATGAGCTTTTACGAAATA	435
Qy	96	AspAsnPheGluLeuTyrPheSerAspAlaPheLeuGlyHisGluIleLeuLysLeu	115
Db	434	AAGAAAAATA---TTTACTTC-----	411
Qy	116	LysAspPhePheAlaTyrIle---LysTyrAsnGluIleTyrAsnArgGlnArgIleThr	134
Db	410	AGAAAATATCTAGAAATAGTTCAAAATAATCCAAACAATCGTTCATTTTAAATGGAATACC	351
Qy	135	Ser-GlyValTyrMetCysAlaThrAlaValAlaLeuGlyTyrIlysserIleTyrIleSe	154
Db	350	TCAAAAATCTGTATCTGTCTTCTAAAAATATCTGT-----ATCTATTCTAA	306
Qy	154	rGlyIleAspPheTyrGlnAspThrAsnLeu-TyrAlaPheAspAsnAsnLysLys-	173
Db	305	AAATATAATAATAATCAATCCATCTTATAGAAAATGGTATTTTAAAAACAAATAATAAAAAA	246
Qy	174	-----AsnLeuLeuAsnLysCys---ThrGlyPheLysAsnGlnL	186
Db	245	ATAAAAATGTTATATAAAACAACAGTAGTAATAAAAAATGTTATTTATCCCAAAAGGCCAGA	186
Qy	186	ysPheLysPheIleAsnHisSerMetAla-----CysAspLeuGlnAlaLeuAspTyrL	204
Db	185	TTTTCCACGTGTATATATATGATATCATCACATTTGTCAATTTGTACCATAGACCACA	126
Qy	204	eumetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyrPheLysLeuA	224
Db	125	TTATAAAAGATATAATATATATATCAATAAATTAATTTAAAC-----	82
Qy	224	laProaspIleGlySerAspPheValLeuSerLysLysProLysLysTyrIleAsnAspI	244
Db	81	-----TTTAAATAAAATAAAAAATGGCTTATTAAACA	51
Qy	244	leLeuIlePro 247	
Db	50	CGATGCTCCCA 40	
RESULT 14			
BJ344875/c			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			

```
/dev stage="Aggregation stage"
/clone_lib="Dictyostelium discoideum cDNA library, Af"

ORIGIN
Alignment Scores:
Pred. No.: 0.137 Length: 699
Score: 103.00 Matches: 63
Percent Similarity: 33.93% Conservative: 32
Best Local Similarity: 22.50% Mismatches: 99
Query Match: 6.56% Indels: 86
DB: 4 Gaps: 15

US-10-735-419-10 (1-294) x BU344875 (1-699)

Qy 12 AsnGlyProSerLeuLysAsnIleAspTyrLysArgLeuProLysGlnPhe---AspVal 30
Db 694 AATTCAAATTCATGAAAGT---AATTTTATTTCATTACCAAGAAAGTTTCAGATTGT 628
Qy 31 PheArgCysAsnGlnPheTyrPheGluAspArgTyrPheValGlyLysAspValLysTyr 50
Db 627 TTT-----ATTATTTTAATGATGATATTTCCTTGGAGCACCAAGTTCGGTG 580
Qy 51 ValPhePheAsnProPheValPhePheGluGlnTyrTyrThrSerLysLysLeuIleGln 70
Db 579 TCAGATTTTGGGATTCAGATGCAGTCAAGCAATTTATAAGAGTAGTTGCACAGCACCA 520
Qy 71 AsnGluGluTyrAsnIleGluAsnIle-----ValCysSerThrIleAsnLeuGluTyr 88
Db 519 CAAAGTAAAGAGAAACAATCGAATATTCGGCATGATGATTTGGCTATPACCAATGATTTA 460
Qy 89 IleAspGlyPheGlnPheValAspAsnPheGluLeuTyrPheSerAspAlaPheLeuGly 108
Db 459 TTGAACAATATTGGGAGGTGATATTAAGAAATCGTCATTAT-----GCATCACATGGT 406
Qy 109 HisGluIleLysLysLeuLysAspPhePheAlaTyrIleLysTyrAsnGluIleTyr 128
Db 405 CATCAA-----TTCTTTTACT-----CGTGAGATCTTT 379
Qy 129 AsnArgGlnArgIleThrSerGlyValTyrMetCysAlaThrAlaValAlaLeuGlyTyr 148
Db 378 GATCGT----- 373
Qy 149 LysSerIleTyrIleSerGlyIleAspPheTyrGlnAspThrAsnAsnLeuTyrAlaPhe 168
Db 372 -----ATGATATCAAGATTGGAAACCAAGATTT----- 346
Qy 169 AspAsnAsnLysLysAsnLeuAsnLysCysThrGlyPheLysAsnGlnLysPheLys 188
Db 345 -----AAGAAACCTCTGCCAATCCA-----TTCCGTACACAGCT----- 313
Qy 189 PheIleAsnHisSerMetAlaCysAspLeuGlnAlaLeuAspTyrLeuMetLysArgTyr 208
Db 312 -----CAGACCTTCAAAATTCCTTTCTTTATTTGGCTTATGTAACTCGTTTCTAT 262
Qy 209 -----AspValAsnIleTyrSerLeuAsnSerAspGluTyrPheLys 222
Db 261 GCTACTTATGAACCAACTCCCAATCAATATTACGCTTTGATCATGATGATCAGAGAA 202
Qy 223 LeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr---Ile 241
Db 201 ATGAAAGAAAGATTGCAAG-----ATTTTAGCAAGAAACCAAGACTGTTGTGTTA 148
Qy 242 AsnAspIleLeuLeuProAspLys----- 249
Db 147 AATGATGTCCTCTCTCTCGATAAACCATAACTAAGTCGTTGATGAATTAAGAGAAATTC 88
Qy 250 -----TyrAlaGlnGlnArgTyrTyrGlyLysLysSerArgLeuLysGluAsnLeu 266
Db 87 TTTGAAACTTATCTTCCAGACGGGTTCATGGGAAATAATCAATAATTAATACTACTCTC 28

RESULT 15
BM162809 787 bp mRNA linear EST 04-DEC-2001
LOCUS
```

DEFINITION EST565332 PyBS Plasmodium yoelii yoelii cDNA clone PYCKX22 5' end, mRNA sequence.

ACCESSION BM162809

VERSION BM162809.1 GI:17308490

KEYWORDS EST.

SOURCE Plasmodium yoelii yoelii

ORGANISM Plasmodium yoelii yoelii

REFERENCE 1 (bases 1 to 787) Eukaryota: Alveolata; Apicomplexa; Haemosporida; Plasmodium.

AUTHORS Carlson, J.M., Daly, T.M., Long, C.A., Bergman, L.W., Vaidya, A.B., Fraser, C.M., and Carucci, D.J.

TITLE Plasmodium yoelii EST project at TIGR

JOURNAL Unpublished (2001)

COMMENT Contact: Jane Carlton
Parasite Genomics Group
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-530-9319
Fax: 301-838-0208
Email: carlton@tigr.org
For clone info, please contact the Malaria Research and Reference Reagent Resource Center, ATCC
http://www.malaria.mr4.org/mr4pages/index.html
Seq primer: ADF.

FEATURES

Location/Qualifiers

1..787

/organism="Plasmodium yoelii yoelii"

/mol_type="mRNA"

/strain="17XL"

/sub_species="yoelii"

/db_xref="taxon:73239"

/clone="PYCKX22"

/dev_stage="Asexual blood stages"

/lab_host="E. coli XL-1 Blue"

/clone_lib="PyBS"

/note="vector: pAD-GAL4; At 20-25% parasitemia, blood was collected from BALB/cByJ mice infected with Py17XL parasites, and leukocytes removed by passage over microcrystalline cellulose columns. Total RNA was isolated using the guanidinium isothiocyanate method, and mRNA isolated using oligo(dT)-cellulose chromatography. First strand cDNA synthesis was completed using a 50-base primer and reverse transcriptase in the presence of 5-methyl dCTP. After second strand synthesis, uneven termini were treated with pfu DNA polymerase and EcoRI adaptors ligated to the blunt ends. The sample was cleaved with XhoI and separated on a Sephacryl S-500 column. Size-fractionated cDNA was precipitated and ligated to HybridZAP arms directionally using EcoRI-XhoI cleaved arms. After packaging, the phagemid vector (pAD-GAL4) was excised from the HybridZAP vector and plasmid DNA isolated."

ORIGIN

Alignment Scores:

Pred. No.: 0.16 Length: 787

Score: 103.00 Matches: 52

Percent Similarity: 39.91% Conservative: 35

Best Local Similarity: 23.85% Mismatches: 109

Query Match: 6.56% Indels: 22

DB: 4 Gaps: 7

US-10-735-419-10 (1-294) x BM162809 (1-787)

Qy 67 LysLeuLeuGlnAsnGluGluTyrAsnIleGluAsnIleValCysSerThrIleAsnLeu 86

Db 43 AAAATATGTTTCAAAAATAATACATATATCAGGAATTTTATAAATTTTCATCTTTAACATA 102

Qy 87 GluTyrIleAspGlyPheGln-----PheValAspAsnPheGluLeuTyr 101

Db 103 GATGATATGATACTATATAAATATATATATATGACAAATTTAATAGAAAATTTGAATCTT 162

Qy 102 PheSerAspAlaPheLeuGlyHisGluIleLysLysLysAspPheAlaTyr 121

```
Db      163  TTATCT-----TTTGAACACAGTAAAAATAAATGAATATAGCTTAAATTTTAAACATCGAA 216
Qy      122  IleLysTyrAsnGluIleTyrAsnArgGlnArgIleThrSerGlyValTyrMetCysAla 141
Db      217  ACATCATATATTAATATATATACAAACAAATAAAAGATTTAATTAATTAATATATGCT 276
Qy      142  ThrAlaValAlaLeuGlyTyrLysSerIleTyrIleSerGlyIleAspPheTyrGlnAsp 161
Db      277  ACATCATATTAGAAAAATATAATACCTTGCATTGAATCAGGACATGAT----- 324
Qy      162  ThrAsnAsnLeuTyrAlaPheAspAsnAsnLysLysAsnLeuLeuAsn---LysCysThr 180
Db      325  -----AATATAGATGCAGAAATACGACAGAAAGAACAAATTAATGAATTCAAATTTATTA 378
Qy      181  GlyPheLysAsnGlnLysPheLysPheIleAsnHisSerMetAlaCysAspLeuGlnAla 200
Db      379  AATTTTATAATAATTTTATTATATAAAAAATCAAAATATT-----CCTAAAAAGTGT 432
Qy      201  LeuAspTyrLeuMetLysArgTyrAspValAsnIleTyrSerLeuAsnSerAspGluTyr 220
Db      433  ATCGAATATAAGATAATATATATGATGAATATAATATAATATAATATAATTTCCAAATTTAT 492
Qy      221  PheLysLeuAlaProAspIleGlySerAspPheValLeuSerLysLysProLysLysTyr 240
Db      493  AATTTTACTGGACTAACACATAAATTAAAGATTTAAACAATCTCTTAGAGTAAAGTTCAA 552
Qy      241  IleAsnAspIleLeuIleProAspLysTyrAlaGlnGluArgTyrTyrGlyLysLysSer 260
Db      553  AATAATAATGAAAAAACTGGAAATTAT---AATACAAAAAATTTATGGAATAAAAAAT 609
Qy      261  Arg-----LeuLysGluAsnLeuHisTyrLysLeuIleLysAsp 273
Db      610  AATTTATTATTTATGAGTTACACAATAATCGGCATGGGGGATTTATTAAGAAT 663
```

Search completed: June 14, 2005, 21:11:37
Job time : 3697 secs

This Page Blank (uspto)